Appunti del corso

Elementi di Analisi e di Ricerca Operativa da un Punto di Vista Superiore

tenuto dal Prof. Paolo Negrini

Università di Bologna - Anno Accademico 2013/2014

Michele Santa Maria

Indice

Introduzione

Questo testo è ricavato dagli appunti da me presi durante il corso *Elementi di Analisi e di Ricerca Operativa da un Punto di Vista Superiore*, tenuto dal professor Paolo Negrini nell'anno accademico 2013/14 all'Università di Bologna.

Nel corso degli appunti ho lasciato come esercizi quelle verifiche che in classe sono state lasciate da fare autonomamente a casa e che risultano prove non troppo complesse o comunque di tipo meccanico. Quegli esercizi che sono stati lasciati e che hanno invece richiesto uno sforzo maggiore li ho riportati come osservazioni dimostrate o esempi.

Prego chiunque legga questo testo di contattarmi nel caso trovasse alcuni errori nel testo (o anche solo per osservazioni/domande/chiarimenti) all'indirizzo email: msm139@gmail.com

Capitolo 1

Un po' di Analisi

In questo capitolo analizzeremo alcuni argomenti un po' casuali riguardanti cose che si incontrano solitamente negli studi scolastici, cercando di capire il perché delle cose e cosa ci sia dietro argomenti anche a prima vista semplici.

1.1 Perché $a^0 = 1$?

24/02/2014

La prima domanda che ci poniamo è proprio questa: Perché $a^0 = 1$?

Una risposta comprensibile anche per studenti di scuola superiore è per far tornare i conti. In effetti la definizione di una potenza nulla è proprio questa per come viene costruita la funzione esponenziale:

Se definiamo l'esponenziale in $\mathbb{N}$ dato dalla semplice moltiplicazione ripetuta

$$n^k := \underbrace{n \cdot n \cdot \ldots \cdot n}_{k \text{ volte}}$$

possiamo dimostrare abbastanza facilmente che vale $n^{h+k} = n^h \cdot n^k$ per ogni $n, h, k \in \mathbb{N}^+$.

Allora per mantenere vera questa proprietà anche nel caso di esponenti nulli (e vogliamo che si mantenga perché $0 \in \mathbb{N}$) dobbiamo *necessariamente* definire $n^0 = 1$ perché abbiamo

$$n^k = n^{k+0} = n^k \cdot n^0$$

Ecco dunque svelato l'arcano.

Per essere più precisi la definizione corretta[1] della funzione esponenziale nell'insieme $\mathbb{N}$ dovrebbe essere di tipo induttivo, ovvero

$$n^k := \begin{cases} 1 & \text{se } k = 0 \\ n \cdot n^{k-1} & \text{altrimenti} \end{cases}$$

in cui definiamo $n^0 = 1$ proprio perché sappiamo dove vogliamo arrivare, cioè al discorso di prima.

1.2 Limiti senza logaritmi

Come potremmo dimostrare che

$$\lim_{n \longmapsto \infty} 2^n = +\infty$$

senza usare la teoria dei logaritmi?

La risposta non è priva di senso, infatti nelle scuole si presenta il logaritmo come la funzione inversa dell'esponenziale, quindi si dovrà trovare un modo di dimostrare questo limite senza utilizzarli!

Per fortuna in questo caso ci aiuta la famosa **Disuguaglianza di Bernulli**

$$(1 + x)^n \geq 1 + nx \qquad \forall n \in \mathbb{N}, \quad \forall x > -1$$

Questa formula si dimostra per induzione su n senza tanti problemi anche per un ragazzo delle superiori (sempre che abbia compreso come fare le dimostrazioni per induzione...), quindi è del tutto lecito utilizzarla.

Grazie a questa il problema è risolto poiché risulta

$$\lim_{n \longmapsto \infty} 2^n \geq \lim_{n \longmapsto \infty} (1 + n) = +\infty$$

1.3 Volume-Superficie di un solido di rotazione

Un qualunque solido ottenuto dalla rotazione lungo l'asse x di una funzione $f(x)$ in un intervallo fissato $[a, b]$ ha un volume calcolabile con l'integrale

$$\int_a^b \pi f^2(x)\, dx$$

e la motivazione di questo è anche facilmente comprensibile: per calcolare il volume non dobbiamo fare altro che integrare l'area dei cerchi con centro x_0 appartenente l'asse x e

[1] Intendiamo "corretta" come "rigorosa", ovvero una definizione coerente con gli assiomi che definiscono $\mathbb{N}$ nella moderna teoria degli insiemi, cosa che ai ragazzi di scuola superiore forse è meglio non sottolineare.

raggio uguale ad $f(x_0)$. Difatti questi cerchi, al variare di x_0 nell'intervallo $[a, b]$, formano proprio il solido di rotazione voluto.

Ma cosa succede se volessimo calcolare la superficie di questo solido?

Una risposta che potrebbe venire abbastanza naturale a questo punto è fare una cosa simile, ma invece che le aree usare i perimetri, dunque calcolare

$$\int_a^b 2\pi f(x)\, dx$$

ma questo è totalmente **sbagliato**!

La formula corretta risulta invece essere

$$\int_a^b \sqrt{1 + f'(x)^2} dx$$

ma allora perché non la si insegna a scuola?

Il problema principale sta nel fatto che di solito nelle scuole si cerca di svolgere argomenti i cui esercizi siano risolvibili in maniera non troppo complessa, ma usando una formula di questo tipo potrebbe succedere di incorrere in problemi la cui soluzione non è affatto semplice.

Ci sono infatti molti esempi di funzioni il cui integrale è difficilmente calcolabile; tra tutti possiamo citare il famoso caso della funzione e^{-x^2} di cui si è dimostrato non esistere una primitiva esprimibile in termini di funzioni elementari.

E questo non è certo l'unico caso, anzi è stato dimostrato che data una famiglia di funzioni qualunque esiste sempre una composizione di queste per cui non è possibile esprimere una primitiva in termini di quelle stesse funzioni.

Dunque il problema è davvero complesso, anche se poteva non sembrare.

1.4 Retta tangente al grafico

28/02/2014

Data la funzione $y = f(x)$, come troviamo la retta tangente al grafico di $f(x)$ nel punto $\big(x_0, f(x_0)\big)$?

In generale con la *derivata* di f; ma si può fare senza derivata? Proviamo a fare alcuni ragionamenti algebrici.

Prendiamo un polinomio $p(x)$.

Quello che dobbiamo fare è cercare un coefficiente angolare m di modo che x_0 risulti soluzione del seguente sistema:

$$\begin{cases} y = p(x_0) + m(x - x_0) \\ y = p(x) \end{cases} \qquad \Longrightarrow \qquad \begin{cases} p(x) - p(x_0) - m(x - x_0) = 0 \\ y = p(x) \end{cases}$$

Teniamo in considerazione la prima equazione, per cercare di ricavare qualcosa su m.

$$p(x) - p(x_0) - m(x - x_0) = 0$$

$$(x - x_0)\left(\frac{p(x) - p(x_0)}{x - x_0} - m\right) = 0$$

$$q(x) := \frac{p(x) - p(x_0)}{x - x_0} \qquad \Longrightarrow \qquad (x - x_0)\big(q(x) - m\big) = 0$$

Il polinomio $q(x)$ appena definito ci dice qualcosa di interessante: x_0 è radice semplice se $m = q(x_0)$, ovvero se

$$m = q(x)\big|_{x=x_0} = \left.\frac{p(x) - p(x_0)}{x - x_0}\right|_{x=x_0}$$

Attenzione: la scrittura $\left.\frac{p(x)-p(x_0)}{x-x_0}\right|_{x=x_0}$ ha senso solo perché sappiamo che quella frazione si può fattorizzare ottenendo il polinomio $q(x)$!

Esempio 1.4.1 $y = p(x) = x^3 - 7x^2 + 20$, $x_0 = 2$

Cerchiamo la tangente nel punto $(x_0, p(x_0)) = (2, 0)$.

$$\begin{cases} y = m(x - 2) \\ y = x^3 - 7x^2 + 20 \end{cases}$$

$$\Longrightarrow \qquad x^3 - 7x^2 + 20 - m(x - 2) = 0$$

$$(x - 2)\left(\frac{x^3 - 7x^2 + 20}{x - 2} - 20\right)$$

$$(x - 2)\big(x^2 - 5x - 10 - m\big) = 0 \qquad \longrightarrow \qquad v(x) := x^2 - 5x - 10 - m$$

Se poniamo $v(x) = 0$ per $x = 2$, otteniamo $m = 4 - 10 - 10 = -24$.

1.5 Numeri Complessi

I *numeri complessi* furono introdotti da **Bombelli**. Si trovano poi nella formula risolutiva delle equazioni di terzo grado di **Cardano**.

Vediamo cosa scriveva Cardano.

Presa l'equazione di terzo grado $y^3 + py + q = 0$, si pone $y = u + v$ così da avere

$$u^3 + v^3 + 3u^2v + 3uv^2 + p(u + v) + q = 0$$

$$u^3 + v^3 + 3uv(u + v) + p(u + v) + q = 0$$

$$u^3 + v^3 + (3uv + p)(u + v) + q = 0$$

Quest'ultima equazione vale, ad esempio, se

$$\begin{cases} u^3 + v^3 = -q \\ uv = -\frac{p}{3} \end{cases}$$

che è un sistema di sesto grado, ma riconducibile ad uno di secondo grado.

$$\begin{cases} u^3 + v^3 = -q \\ (uv)^3 = \left(-\frac{p}{3}\right)^3 \end{cases}$$

In questo cado le soluzioni u^3, v^3 si possono trovare risolvendo la seguente equazione:

$$t^2 + qt - \left(\frac{p}{3}\right)^3 = 0$$

$$\Longrightarrow \qquad t = \frac{-q \pm \sqrt{q^2 + 4\left(\frac{p}{3}\right)^3}}{2} = -\frac{q}{2} \pm \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}$$

Quindi, ricordando che $y = u + v = \sqrt[3]{u^3} + \sqrt[3]{v^3}$, otteniamo:

$$y = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3}}$$

che è la famosa **Formula di Cardano**.

Esempio 1.5.1 $y^3 - 15y - 126 = 0$

$$p = -15, \quad q = -126$$

$$\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3 = 63^2 - 125 = 62^2$$

$$y = \sqrt[3]{63 + 62} + \sqrt[3]{63 - 62} = 5 + 1 = 6$$

$$\Longrightarrow \qquad y^3 - 15y - 126 = (y - 6)(y^2 + 6y + 21)$$

Esempio 1.5.2 $y^3 - 15y - 4 = 0$

$$p = -15, \quad q = -4$$

$$\left(\frac{q}{2}\right)^2 - \left(\frac{p}{3}\right)^3 = 4 - 125 = -121$$

In questo caso la formula di Cardano non funziona perché ho $\sqrt{-121}$ che non sta in $\mathbb{R}$.

Se invece consideriamo anche i numeri complessi potremmo comunque proseguire e trovare

$$y = \sqrt[3]{2 + 11i} + \sqrt[3]{2 - 11i}$$

1.6 Ricerca della primitiva giusta

Poniamo di voler risolvere il seguente integrale:

$$\int_0^{\frac{\pi}{2}} \frac{1}{5 + 3\sin x}\, dx$$

che potrebbe venir fuori, ad esempio, dallo studio di un moto armonico.

Come facciamo?

Presto detto: abbiamo già fatto tante volte la sostituzione $t = \tan\left(\frac{x}{2}\right)$ che risolve questo tipo di problemi, quindi applichiamola al nostro caso.

$$t = \tan\left(\frac{x}{2}\right)$$

$$\sin x = \frac{2t}{1+t^2} \qquad \cos x = \frac{1-t^2}{1+t^2}$$

$$dx = \frac{2}{1+t^2}\, dt \qquad dt = \frac{1 + \tan\left(\frac{x}{2}\right)}{2}\, dx$$

$$\int_0^{\frac{\pi}{2}} \frac{1}{5 + 3\sin x}\, dx = \int_0^1 \frac{1}{5 + \frac{6t}{1+t^2}} + \frac{2}{1+t^2}\, dt = \int_0^1 \frac{2}{5 + 5t^2 + 6t}\, dt = \otimes$$

Ricordando che quando $\Delta < 0$ vale

$$\int \frac{1}{ax^2 + bx + c}\, dx = \frac{2}{\sqrt{-\Delta}} \arctan\left(\frac{2ax + b}{\sqrt{-\Delta}}\right)$$

risolviamo il nostro integrale:

$$\otimes = \left[2 \cdot \frac{2}{64} \arctan\left(\frac{10t+6}{\sqrt{64}}\right)\right]_0^1 = \frac{1}{2}\left[\arctan\left(\frac{5t+3}{4}\right)\right]_0^1 =$$
$$\frac{1}{2}\left[\arctan 2 - \arctan\left(\frac{3}{4}\right)\right]$$

Possiamo ora ricondurci alla sostituzione iniziale per trovare una primitiva della funzione di partenza!

$$\frac{1}{2}\left[\arctan 2 - \arctan\left(\frac{3}{4}\right)\right] = \frac{1}{2}\left[\arctan\left(\frac{5\tan\frac{x}{2}+3}{4}\right)\right]_{x=0}^{x=\frac{\pi}{2}} = 0$$

da cui

$$\int_0^{2\pi} \frac{1}{5+3\sin x}\,dx = \frac{1}{2}\left[\arctan\left(\frac{5\tan\frac{x}{2}+3}{4}\right)\right]_{x=0}^{x=2\pi} = 0$$

Ma com'è possibile che il risultato sia *zero*?

Il denominatore non si annulla mai, visto che 5 maggiora sicuramente $e \sin x$. Ma allora cos'è andato storto?

Quello che non va bene è la primitiva. Ovvero, avevamo un'espressione che ha come derivata la funzione integranda, che però on va bene per alcuni valori. Questo perché la sostituzione di partenza ($t = \tan\left(\frac{x}{2}\right)$) non va bene per $x = \pi + 2k\pi$.

Dobbiamo aggiustare le cose, cioè ridefinire la primitiva anche nei punti incriminati:

$$\Psi(x) = \begin{cases} \frac{1}{2}\arctan\left(\frac{5\tan\left(\frac{x}{2}\right)+3}{4}\right) & \text{se } x \in]-\pi, \pi[\\ \frac{\pi}{4} & \text{se } x = \pi \\ \frac{1}{2}\arctan\left(\frac{5\tan\left(\frac{x}{2}\right)+3}{4}\right) + \frac{\pi}{2} & \text{se } x \in]\pi, 3\pi[\end{cases}$$

e allora con questa primitiva il calcolo dell'integrale di partenza diventa semplicemente

$$\int_0^{2\pi} \frac{1}{5+3\sin x}\,dx = \Psi(2\pi) - \Psi(0)$$

1.7 Celebri esempi di norme

13/03/2014

In $\mathbb{R}^n$ possiamo definire diverse norme; le più classiche sono

$$\|x\|_\infty := \max |x_k| \qquad \|x\|_1 := \sum_{k=1}^{n} x_k$$

$$\|x\|_2 := \sqrt{\sum_{k=1}^{n} x_k^2} \qquad \|x\|_p := \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}$$

Esercizio 1.7.1 Chi sono gli insiemi $\{x \in \mathbb{R}^n \mid \|x\| = 1\}$ al variare delle diverse norme?

In $\mathbb{R}^n$ a partire da un prodotto scalare del tipo

$$< x, y >:= \sum x_k y_k$$

è possibile definire la norma

$$\|x\|_2 := \sqrt{< x, y >}$$

che induce una funzione distanza

$$d(x, y) := \|x - y\|$$

così da arrivare a scoprire che $\mathbb{R}^n$ è metrico!

Possiamo anche definire delle norme sullo spazio delle funzioni continue nell'intervallo $[a, b]$ (indicate con $f \in \mathcal{C}([a, b])$):

$$\|f\|_\infty := \max |f(x)| \qquad \|f\|_1 := \int_a^b |f(x)| dx$$

$$\|f\|_2 := \sqrt{\int_a^b (f(x))^2 \, dx} \qquad \|f\|_p := \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$$

Allo stesso modo è possibile fare lo stesso per funzioni $\mathcal{C}^1([a, b])$:

$$\|f\|_{1,\infty} := \max |f| + \max |f'| \qquad \|f\|_{1,1} := \int_a^b \left(|f| + |f'|\right)$$

$$\|f\|_{1,2} := \sqrt{\int_a^b (|f|^2 + |f'|^2)} \qquad \|f\|_{1,p} := \left(\int_a^b \left(|f|^p + |f'|^p\right)\right)^{\frac{1}{p}}$$

Lo spazio $\mathcal{C}^1([a, b])$ risulta completo con la norma $\| \cdot \|_{1,\infty}$, ma con tutte le altre no, il che rende problematico ragionare sulle successioni convergenti in questo spazio se non usiamo proprio questa norma.

Quello che possiamo fare per poter parlare di successioni anche con le altre norme è considerare il *completamento* del nostro spazio, proprio come è stato fatto per costruire $\mathbb{R}$ a partire da $\mathbb{Q}$, ricordiamolo:

Definiamo sullo spazio $\mathcal{X} = \{q_n \text{ successione di Cauchy in } \mathbb{Q}\}$ la relazione di equivalenza

$$q_n \sim q'_n \iff \lim_{n\to+\infty} (q_n - q'_n) = 0$$

e prendiamo $\mathbb{R} := \mathcal{X}/_{\sim}$.

Allo stesso identico modo facciamo con lo spazio $\mathcal{C}([a,b])$:
Preso $\mathcal{Y} = \big\{f_n$ successione di Cauchy in $\big(\mathcal{C}([a,b]), \|\cdot\|_1\big)\big\}$ la relazione di equivalenza

$$f_n \sim g_n \quad \Longleftrightarrow \quad \lim_{n\to+\infty}\left(\int_a^b |f_n - g_n|\right) = 0$$

e prendiamo $L^1([a,b]) := \mathcal{Y}/_{\sim}$.

O anche con lo spazio $\mathcal{C}^1([a,b])$:
Preso $\mathcal{Z} = \{f_n$ successione di Cauchy in $\big(\mathcal{C}^1([a,b]), \|\cdot\|_{1,1}\big)\}$ la relazione di equivalenza

$$f_n \sim g_n \quad \Longleftrightarrow \quad \lim_{n\to+\infty}\left(\int_a^b \Big(|f_n - g_n| + |f_n' - g_n'|\Big)\right) = 0$$

e prendiamo $H^{1,1}([a,b]) := \mathcal{Z}/_{\sim}$.

Nota: questa costruzione è equivalente a quella fatta con la *derivata debole*, anche se spesso non viene presentata nei corsi di analisi dei primi anni di università.

Definizione 1.7.1 Presa $f : [a,b] \longrightarrow \mathbb{R}$ diciamo che $g : [a,b] \longrightarrow \mathbb{R}$ è la *derivata debole* di f se per ogni $\varphi \in \mathcal{C}_0^1([a,b])$ vale

$$\int_a^b f \cdot \varphi' = -\int_a^b g \cdot \varphi$$

Capitolo 2

Esponenziali, Logaritmi e funzioni circolari

2.1 Logaritmi

04/03/2014

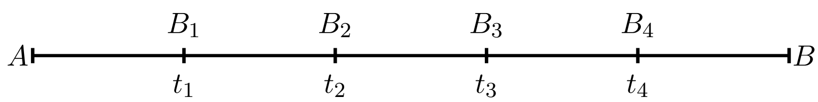
Una tavola dei Logaritmi non è altro che una corrispondenza tra una progressione geometrica ed una aritmetica, ad esempio

n	1	2	4	8	16	32
$\log_2 n$	0	1	2	3	4	5

Una progressione aritmetica possiamo anche vederla come rappresentazione di un modo costante su una retta.

Ad esempio, supponiamo che una particella parta da un punto A della retta, per arrivare ad un altro punto B della retta. Se iniziamo a misurare la sua posizione ad intervalli di tempo costanti, troveremo che all'istante t_1 essa si trova nel punto B_1, che ha una certa distanza da A: $d(A, B_1)$. All'istante t_2 si trova nel punto B_2 che ha distanza da B_1 esattamente la stessa: $d(B_1, B_2) = d(A, B_1)$. E così via fino ad arrivare a B.

Possiamo rappresentare facilmente la situazione in questo modo

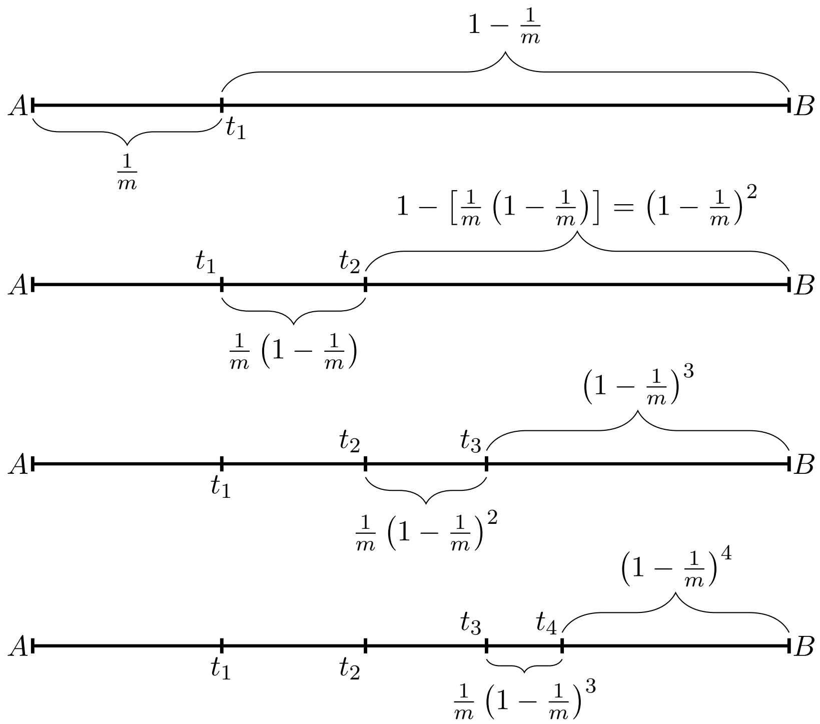


In questa visione la progressione geometrica a cosa corrisponde?

Corrisponde ad un moto che ha uno strano andamento. Scelto un numero $m \in \mathbb{N}^+$ (che sarà la *base* del nostro logaritmo), misuriamo la sua posizione ad intervalli di tempo costanti, come facevamo prima. Stavolta però il punto parte da A e al tempo t_1 ha percorso

$\frac{1}{m}$-esimo del tragitto che lo separa da B. All'istante t_2 di nuovo ha percorso $\frac{1}{m}$-esimo del tragitto che lo separa da B, e così anche all'istante t_3, e così via.

La situazione è ben più complessa da rappresentare! Supponiamo, per semplicità, che la distanza $d(A, B)$ sia uguale a 1. Allora quello che viene fuori si riassume nei seguenti disegni:



E così via.

Quello che succede è che la particella...rallenta!

Ma questo moto ha senso? Esiste nel mondo reale?

Ebbene si, è proprio il moto che compie un corpo che rotola su un piano sotto la forza di gravità, e può essere espresso con il seguente problema di Cauchy:

$$\begin{cases} y'' = 0 \\ y(0) = 0 \\ y'(0) = v_0 \end{cases}$$

Questo moto può anche essere studiato rispetto ad una situazione leggermente differente. Pensiamo un ciclista che si muove in assenza di vento globale, quindi che si trova contro un vento proporzionale alla velocità con cui si muove. Le equazioni diventano queste:

$$\begin{cases} y'' = -py' \\ y(0) = 0 \\ y'(0) = v_0 \end{cases}$$

e la soluzione sappiamo trovarla

$$\begin{cases} y(t) = c_1 + c_2 e^{-pt} \\ y'(t) = -pc_2 e^{-pt} \\ y(0) = 0 & \longrightarrow & c_1 + c_2 = 0 \\ y'(0) = v_0 & \longrightarrow & -pc_2 = v_0 \end{cases}$$

$$\Longrightarrow \quad c_2 = -c_1 = -\frac{v_0}{p}$$

$$y(t) = \frac{v_0}{p}\left(1 - e^{-pt}\right)$$

Quello che abbiamo trovato rappresenta ancora una successione geometrica perché il rapporto tra gli spazi percorsi in un lasso di tempo fissato è costante! Infatti

$$\frac{y\big((k+1)t\big) - y(kt)}{y(kt) - y\big((k-1)t\big)} = e^{-pt}$$

che non dipende da k, quindi resta costante.

Seguendo questo ordine di idee, **Nepero** (1550-1617) presenta il *logaritmo* in questo modo:

Si assume la misura di $\overline{AB_n}$ come il logaritmo di $\overline{t_nB}$.

$$\log\left(\left(1 - \frac{1}{m}\right)^n\right) = \frac{n}{m}$$

Nello specifico, Nepero sceglie $m = 10^7$.

Potremmo chiederci una cosa, visto che non viene specificata: ma che base ha questo logaritmo?

Se $n = m$ abbiamo

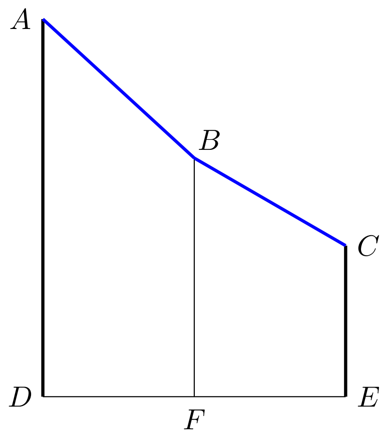
$$\log\left(\left(1 - \frac{1}{m}\right)^n\right) = 1 \qquad \Longrightarrow \qquad b = \left(1 - \frac{1}{m}\right)^m \cong \frac{1}{e}$$

in cui e è proprio la *e di Nepero*!

2.1.1 Logaritmi alla Torricelli

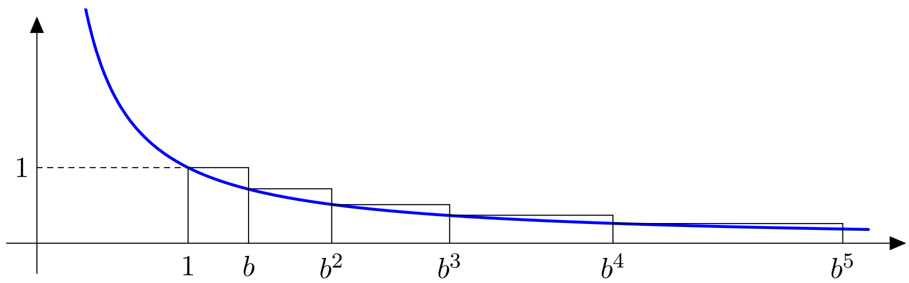
Anche **Torricelli** (1608-1647) studiò i logaritmi, ma in modo differente rispetto a Nepero.

Dati due segmenti, AD e CE, prendiamo F il punto medio di DE, e costruiamo il segmento FB in modo che sia la media geometrica tra AD e CE.



Il procedimento può essere iterato per formare una curva logaritmica.

Torricelli trovò un logaritmo anche integrando un ramo di iperbole nel seguente modo[1]



Area rettangolo di base $\overline{1\,b} = b - 1$.
Area rettangolo di base $\overline{b\,b^2} = (b^2 - b)b^{-1} = b - 1$.
$\vdots$
Area rettangolo di base $\overline{b^n\,b^{n+1}} = b - 1$.

Quindi di nuovo abbiamo la progressione geometrica delle basi dei rettangoli considerati.

$$n(b-1) = \log\left(b^n\right)$$

[1] Nel grafico gli assi sono riscalati per capire meglio il disegno.

in cui stavolta la base è

$$b = \left(1 + \frac{1}{n}\right)^n \quad \xrightarrow{n\to\infty} \quad e$$

da cui abbiamo la seguente integrazione:

$$\int_1^c \frac{1}{x}\,dx = \log_e c$$

2.2 Esponenziale complesso e funzioni circolari

07/03/2014

Un modo interessante di presentare la funzione esponenziale è quello di partire direttamente col definire l'esponenziale complesso nel seguente modo:

Per ogni $z \in \mathbb{C}$ definiamo

$$e^z = \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Essendo la serie assolutamente convergente la definizione è ben posta, e la funzione $\exp(z)$ (o e^z) risulta olomorfa su tutto $\mathbb{C}$.

Vediamone le proprietà principali:

1. $\exp(a+b) = \exp(a)\cdot\exp(b)$
2. $\exp(z) \neq 0 \ \ \forall z \in \mathbb{C}$
3. $\exp'(z) = \exp(z)$
4. $\exp(z)\big|_{\mathbb{R}}$ è strettamente crescente e valgono
$$\lim_{x\to+\infty}\exp(x) = +\infty \qquad \lim_{x\to-\infty}\exp(x) = 0$$
5. $\exp(z + 2k\pi i) = \exp(z) \ \ \forall z \in \mathbb{C},\ \forall k \in \mathbb{Z}$
6. $\exp(it)$, con $t \in \mathbb{R}$, applica $\mathbb{R}$ su $\gamma := \{z \in \mathbb{C} \mid |z| = 1\}$
7. $\exp(z)$ è suriettiva su $\mathbb{C}\setminus\{0\}$
8. $\exists\, \pi \in \mathbb{R}$ tale che $\exp(i\pi) + 1 = 0$

A partire dall'esponenziale complesso è possibile definire (o trovare, a seconda dei casi) le funzioni circolari "seno" e "coseno".

Se consideriamo $z = it \in \mathbb{C}$ possiamo riscrivere l'esponenziale complesso come

$$e^{it} = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{(2n)!} + i \cdot \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} := \cos(t) + i\sin(t)$$

ed avere dunque una formulazione delle funzioni circolari come somme infinite.

Anche se definite in maniera diversa dal solito, per queste funzioni si possono ritrovare le proprietà fondamentali, ad esempio

Proposizione 2.2.1 *Valgono le seguenti uguaglianze:*

$$\sin'(t) = \cos(t)$$
$$\cos'(t) = -\sin(t)$$

Dimostrazione.

$$\frac{d}{dt} e^{it} = ie^{it} = i(\cos(t) + i\sin(t)) = -\sin(t) + i\cos(t)$$

$$\frac{d}{dt} e^{it} = \frac{d}{dt}(\cos(t) + i\sin(t)) = \cos'(t) + i\sin'(t)$$

$$\implies \quad \begin{cases} \sin'(t) = \cos(t) \\ \cos'(t) = -\sin(t) \end{cases}$$

□

Proposizione 2.2.2 *Per ogni $t \in \mathbb{R}$ vale la seguente equazione:*

$$\cos^2(t) + \sin^2(t) = 1$$

Dimostrazione.

$$1 = e^{it} \cdot e^{-it} = e^{it} \cdot \overline{e^{it}} = \left|e^{it}\right|^2 = \cos^2(t) + \sin^2(t)$$

□

Osservazione $\exists\, t > 0$ tale che $\cos(t) = 0$.

Dimostrazione. Per dimostrarne l'esistenza possiamo osservare che $\cos(0) = 1$ e $\cos(2) < 0$ perché:

$$\cos(2) = \sum_{n=0}^{\infty} (-1)^n \frac{2^{2n}}{(2n)!} = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} + \frac{2^8}{8!} - \ldots =$$

$$= 1 + \underbrace{\frac{-2^2 \cdot 3 \cdot 4 + 2^2 \cdot 2 \cdot 2}{4!}}_{<-1} + \underbrace{\frac{-2^6 \cdot 7 \cdot 8 + 2^6 \cdot 2 \cdot 2}{4!}}_{<0} + \ldots < 0$$

Dunque per continuità della funzione $\cos(t)$ deve esistere un $\widehat{t} \in]0, 2[$ tale che $\cos\left(\widehat{t}\right) = 0$.

□

A questo punto possiamo definire

$$\frac{\pi}{2} := \min\{t \in \mathbb{R}^+ \mid \cos(t) = 0\}$$

che esiste poiché l'insieme è non vuoto e compatto[2] in $\mathbb{R}^+ \cap [0, 2]$.

Da questa definizione e dalla relazione $\cos^2(t) + \sin^2(t) = 1$ segue immediatamente che $\sin(\frac{\pi}{2})$ deve essere $+1$ o -1.

Ma poiché $\sin'(t) = \cos(t)$ il seno cresce nell'intervallo $[0, \frac{\pi}{2}[$, dunque risulta

$$\sin\left(\frac{\pi}{2}\right) = 1$$

Proposizione 2.2.3

$$e^z = 1 \iff z = 2k\pi i \ \textit{con} \ k \in \mathbb{Z}$$

Dimostrazione. $\boxed{\Leftarrow}$ Da ciò che abbiamo trovato fin'ora possiamo subito dire che $e^{\frac{\pi}{2}i} = i$, da cui otteniamo $e^{i\pi} = -1$, ovvero $e^{2\pi i} = 1$.

Da quest'ultima relazione viene immediatamente una delle proprietà elencate sopra

$$e^{2k\pi i} = \left(e^{2\pi i}\right)^k = 1^k = 1 \ \ \forall k \in \mathbb{Z}$$

$\boxed{\Rightarrow}$ Sapendo che $e^z = 1$ vediamo cosa possiamo ricavare su $z = x + iy$.

$$|e^z| = 1 \quad \Longrightarrow \quad |e^{x+iy}| = |e^x(\cos(y) + i\sin(y))| = 1$$

$$\Longrightarrow \quad e^x = 1 \quad \Longrightarrow \quad x = 0$$

Quindi la parte reale di z è nulla.

Sapendo che la funzione e^{iy} è periodica, che abbiamo dimostrato prima, ci basterà dimostrare che $\forall y \in]0, 2\pi[$ vale $e^{iy} \neq 1$.

Supponiamo per assurdo che $\exists\, w \in]0, 2\pi[$ tale che $e^{iw} = 1$.

Da ciò ci scriviamo $e^{i\frac{w}{4}}$ come un certo numero complesso $u + vi$, ovvero

$$\cos\left(\frac{w}{4}\right) + i\sin\left(\frac{w}{4}\right) = e^{i\frac{w}{4}} = u + vi$$

e ricaviamo

$$1 = e^{iw} = \left(e^{i\frac{w}{4}}\right)^4 = (u + vi)^4 = (u^2 + 2uvi - v^2)^2 =$$

$$= u^4 + 4u^3vi - 4u^2v^2 - 4uv^3i - 2u^2v^2 + v^4 = (u^4 - 6u^2v^2 + v^4) + i(4u^3v - 4uv^3)$$

$$\Longrightarrow 4u^3v - 4uv^3 = 0 \quad \Longleftrightarrow \quad uv(u^2 - v^2) = 0$$

Ma $u = \cos(\frac{w}{4}) > 0$, $v = \sin(\frac{w}{4}) > 0$ quindi vale $u^2 = v^2 = \frac{1}{2}$.

[2]Ricordiamo che compatto in $\mathbb{R}^n$ equivale a chiuso e limitato, dunque nel nostro caso la chiusura è garantita dal fatto che l'insieme preso è controimmagine di un punto di una funzione continua.

Riportando questi risultati nelle equazioni precedenti otteniamo

$$1 = e^{iw} = \frac{1}{4} + \frac{1}{4} - \frac{3}{2} = -1$$

che è assurdo.

$\square$

Con questa introduzione alle funzioni circolari risultano molto più semplici da dimostrare le formule di somma, basta infatti applicare le proprietà appena dimostrate dell'esponenziale complesso ed ottenere

$$\cos(\alpha+\beta) + i\sin(\alpha+\beta) = e^{i(\alpha+\beta)} = e^{i\alpha}\cdot e^{i\beta} =$$

$$= \Big(\cos(\alpha) + i\sin(\alpha)\Big)\Big(\cos(\beta) + i\sin(\beta)\Big) =$$

$$= \Big(\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)\Big) + i\Big(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)\Big)$$

$$\Longrightarrow \quad \begin{cases} \cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \sin(\alpha+\beta) = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) \end{cases}$$

Proposizione 2.2.4 $\exp(it)$, *con* $t \in \mathbb{R}$, *applica* $\mathbb{R}$ *su* $\gamma := \{z \in \mathbb{C} \mid |z| = 1\}$.

Dimostrazione. Dobbiamo dimostrare che e^{it} è suriettiva su γ quando $t \in \mathbb{R}$.

Preso $z = u + vi \in \gamma$, sicuramente $u, v \in]0,1[$. Concentriamoci inizialmente sul primo quadrante del piano complesso.

Visto che cerchiamo un $t \in \mathbb{R}$ tale che $e^{it} = z$ ragioniamo sull'equazione

$$u + vi = z = e^{it} = \cos(t) + i\sin(t)$$

Nel primo quadrante vale $\sin(0) = 0$, $\sin(\frac{\pi}{2}) = 1$ quindi $\exists!\, \tau \in]0, \frac{\pi}{2}[$ tale che $\sin(\tau) = v$, da cui

$$u^2 = 1 - v^2 = 1 - \sin^2(\tau) = \cos^2(\tau)$$

Dunque $\cos(\tau)$ è uguale a u o $-u$. Ma u è positivo, e $\cos(\tau)$ è positivo nel primo quadrante, quindi deve essere $\cos(\tau) = u$.

$\Longrightarrow \quad e^{i\tau} = u + vi = z$

Nel caso fossimo negli altri quadranti possiamo ragionare in modo sostanzialmente uguale giocando con la positività delle funzioni seno e coseno e trovare lo stesso risultato.

$\square$

Proposizione 2.2.5 $\exp(z)$ *è suriettiva su* $\mathbb{C} \setminus \{0\}$.

Dimostrazione. Preso $\omega \in \mathbb{C} \setminus \{0\}$ possiamo scrivere

$$\omega = |\omega| \frac{\omega}{|\omega|}$$

in cui $\frac{\omega}{|\omega|} \in \gamma$ definita sopra.

Ma allora la proposizione è un corollario della precedente, poiché vale

$$e^z = e^{x+iy} = e^x \cdot e^{iy}$$

e sappiamo già che e^{iy} è suriettiva su γ, quindi basterà scegliere y tale che $e^{iy} = \frac{\omega}{|\omega|}$ e x tale che $e^x = |\omega|$, che sappiamo di poter fare perché $\exp(\cdot)\big|_{\mathbb{R}}$ è suriettiva su $]0, +\infty[$. $\square$

Quest'ultima proposizione ci suggerisce anche un nuovo modo per rappresentare un numero complesso:

$$e^z = \omega \quad \Longleftrightarrow \quad z = \ln|\omega| + i \cdot arg(\omega)$$

in cui la funzione $arg(\omega)$ è definita nel seguente modo

$$arg(\omega) = t \quad \Longleftrightarrow \quad \frac{\omega}{|\omega|} = \cos(t) + i\sin(t)$$

Possiamo anche decidere di usare un simbolo per definire l'inversa dell'esponenziale complesso

$$Log(\omega) := \ln|\omega| + i \cdot arg^*(\omega)$$

in cui la funzione $arg^*(\omega)$ è semplicemente $arg(\omega)$ ristretta a $[0, 2\pi[$.

C'è un problema però con questa funzione: non è continua!

Infatti $Log(1) = 0$, e preso un qualsiasi $\alpha \in [0, 2\pi[$ vale $Log(\cos(\alpha) + i\sin(\alpha)) = i\,\alpha$, ma purtroppo

$$\lim_{\alpha \to 2\pi} Log(\cos(\alpha) + i\sin(\alpha)) = 1 \neq 0$$

2.3 Problemi di Cauchy

2.3.1 Funzioni circolari

Dato il seguente problema di Cauchy:

$$\begin{cases} f' = g \\ g' = -f \\ f(0) = 0 \\ g(0) = 1 \end{cases}$$

Possiamo *definire* le soluzioni come

$$\sin(x) := f(x)$$
$$\cos(x) := g(x)$$

Una volta definite in questo modo le funzioni seno e coseno possiamo studiarne le proprietà per "scoprire" che sono proprio le funzioni a cui siamo solitamente abituati!

Proposizione 2.3.1 *$f(x) = \sin(x)$ è dispari; $g(x) = \cos(x)$ è pari.*

Dimostrazione. Definiamo due funzioni ausiliarie:
$f_1(x) := -f(-x)$
$g_1(x) := g(-x)$

E calcoliamo le loro derivate:

$$f_1'(x) = -f'(-x) \cdot (-1) = f'(-x) = g(-x) = g_1(x)$$

$$g_1'(x) = g(-x) \cdot (-1) = -g'(-x) = f(-x) = -f_1(x)$$

Dunque per f_1 e g_1 vale

$$\begin{cases} f_1' = g_1 \\ g_1' = -f_1 \end{cases}$$

Quindi per unicità della soluzione del problema di Cauchy risulta $f = f_1$ e $g = g_1$.

□

Proposizione 2.3.2 *Per ogni $x \in \mathbb{R}$ vale la seguente relazione*

$$f^2(x) + g^2(x) = \sin^2(x) + \cos^2(x) = 1$$

Dimostrazione. Basta osservare che la funzione $f^2 + g^2$ è costante, difatti

$$\left(f^2 + g^2\right)' = 2ff' + 2gg' = 2fg - 2gf = 0$$

da cui otteniamo che la funzione assume un solo valore, che possiamo trovare calcolando la funzione in un punto qualsiasi, ad esempio lo zero:

$$f^2(x) + g^2(x) = f^2(0) + g^2(0) = 1 \qquad \forall x$$

□

Corollario *Per ogni $x \in \mathbb{R}$ valgono le seguenti relazioni:*

$$-1 \leq \cos(x) \leq 1 \qquad\qquad -1 \leq \sin(x) \leq 1$$

Proviamo a calcolare le **Serie di Taylor** delle nostre nuove funzioni:

$$f(x) = \sum_{n=0}^{\infty} a_n \, x^{2n+1}$$

$$f'(x) = \sum_{n=0}^{\infty} (2n+1) a_n \, x^{2n}$$

$$f''(x) = \sum_{n=0}^{\infty} (2n+1)(2n) a_n \, x^{2n-1} = \sum_{m=0}^{\infty} (2m+3)(2m+2) a_{m+1} \, x^{2m+1}$$

in cui l'ultima uguaglianza è solamente un cambio di indici.

Ma ragionando sulla definizione delle nostre funzioni f e g di partenza ricordiamo che vale $f' = g$, $g' = -f \implies f'' = -f \implies f'' + f = 0$. Se usiamo questa relazione con le serie definite sopra otteniamo

$$0 = f''(x) + f(x) = \sum_{n=0}^{\infty} \big(a_n + (2n+3)(2n+2) a_{n+1}\big) x^{2n+1}$$

$$\implies \qquad a_{n+1} = \frac{-a_n}{(2n+3)(2n+2)}$$

Quindi abbiamo trovato una formulazione implicita dei coefficienti della serie di Taylor, che adesso possiamo calcolare esplicitamente:

$a_0 = f'(0) = 0$
$a_1 = \frac{-1}{3 \cdot 2} = -\frac{1}{3!}$
$a_2 = \frac{1}{3! \cdot 4 \cdot 5} = \frac{1}{5!}$
$\vdots$
$a_n = \frac{(-1)^n}{(2n+1)!}$

Ed ecco trovata la serie di Taylor:

$$\sin(x) = f(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

In maniera del tutto analoga possiamo trovare la serie del coseno:

$$\cos(x) = g(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

10/03/2014

Proposizione 2.3.3 (Formule di Addizione) $\forall x \in \mathbb{R}$ *valgono le seguenti relazioni:*

$$\begin{cases} f(x+y) = f(x)g(y) + f(y)g(x) \\ g(x+y) = g(x)g(y) - f(x) + f(y) \end{cases}$$

Dimostrazione. Definiamo le funzioni $f_1(x) = f(x+y)$ e $g_1(x) = g(x+y)$, ovvero pensiamo la y fissata e facciamo variare la x.

f_1 e g_1 soddisfano lo stesso sistema di equazioni differenziali che definisce le funzioni f e g, però non soddisfano le soluzioni iniziali, ovvero vale

$$\begin{cases} f_1' = g_1 \\ g_1' = -f_1 \end{cases}$$

Proviamo a fare un po' di calcoli con queste funzioni:

$$(fg_1 - f_1g)' = f'g_1 + fg_1' - f_1'g - f_1g' = gg_1 - ff_1 - g_1g + f_1f = 0$$

Dunque la funzione è costante: $f(x)g_1(x) - f_1(x)g(x) = a \in \mathbb{R} \quad \forall x$.
Allo stesso modo calcoliamo

$$(ff_1 + gg_1)' = f'f_1' + ff_1' + g'g_1 + gg_1' = gf_1 + fg_1 - fg_1 - gf_1 = 0$$

Quindi anche questa risulta costante: $f(x)f_1(x) + g(x)g_1(x) = b \in \mathbb{R} \quad \forall x$.
Calcoliamo ora

$$af + bg = (f^2g_1 - f_1gf) + (ff_1g + g^2g_1) = g_1(f^2 + g^2) = g_1$$

$$ag - bf = (fg_1g - f_1g^2) - (f^2f_1 + gg_1f) = -f_1$$

Per trovare il valore delle costanti a e b ragioniamo così: se poniamo $x = 0$ otteniamo

$$a = f(0)g_1(0+y) - f_1(0+y)g(0) = -f(y)$$

$$b = f(0)f_1(0+y) + g(0)g_1(0+y) = g(y)$$

Concludendo così la nostra dimostrazione:

$$g(x+y) = g_1(x) = af(x) + bg(x) = -f(y)f(x) + g(y)g(x)$$

$$f(x+y) = f_1(x) = ag(x) - bf(x) = f(y)g(x) + g(y)f(x)$$

□

Teorema 2.3.1 $\exists\, \widehat{x} > 0$ *tale che* $\sin(\widehat{x}) = 1$.

Dimostrazione. Sappiamo già che $-1 \leq \sin(x) \leq 1 \;\; \forall x$.

Supponiamo per assurdo che si abbia $-1 < \sin(x) < 1 \;\; \forall x > 0$. Allora si deve avere $|\cos(x)| > 0 \; \forall x > 0$.

Ma visto che $\cos(0) = 1 > 0 \Longrightarrow \sin'(x) = \cos(x) > 0 \;\; \forall x > 0$.

Allora $\sin(x)$ è strettamente crescente in $\mathbb{R}^+$, e sappiamo che $\cos'(x) = -\sin(x) < 0$, quindi $\cos(x)$ è strettamente decrescente in $\mathbb{R}^+$.

Prendiamo dunque un positivo fissato $a > 0$ e dimostriamo per induzione che $\cos(2^n a) \leq (\cos(a))^{2^n}$.

$\boxed{n = 1}$

$$\cos(2a) = (\cos(a))^2 - (\sin(a))^2 \leq (\cos(a))^2$$

$\boxed{n \Rightarrow n+1}$

$$\cos(2^{n+1}a) = \cos(2 \cdot 2^n a) = (\cos(2^n a))^2 - (\sin(2^n a))^2 \leq (\cos(2^n a))^2 \leq (\cos(a)^{2^n})^2 = (\cos(a))^{2^{n+1}}$$

Avendo dimostrato questa proprietà adesso sappiamo che

$$\lim_{n \longrightarrow +\infty} \cos(2^n a) = 0 \quad \Longrightarrow \quad \exists\, b > 0 \text{ tale che } \cos(b) < \frac{1}{4}$$

$$\Longrightarrow \quad \begin{cases} \cos^2(b) < \frac{1}{16} \\ \sin^2(b) > \frac{15}{16} \end{cases}$$

Dunque

$$\cos(2b) = \cos^2(b) - \sin^2(b) < \frac{1}{16} - \frac{15}{16} = -\frac{14}{16} < 0$$

il che ci porta ad un assurdo perché il coseno doveva essere sempre positivo.

$\square$

Definizione 2.3.1 Avendo dimostrato che $\exists\, x > 0$ tale che $\sin(x) = 0$ definiamo il seguente simbolo:

$$\frac{\pi}{2} := \min\{x \geq 0 \mid \sin(x) = 1\}$$

Corollario $\cos\left(\frac{\pi}{2}\right) = 0$

Riporto un secondo modo di definire $\frac{\pi}{2}$, non svolto a lezione, ma che mi sembra più chiaro:

Teorema 2.3.2 *$\exists\, \widehat{x} > 0$ tale che $\cos(\widehat{x}) = 0$.*

Dimostrazione. Supponiamo per assurdo che un tale $\widehat{x}$ non esista. Allora $\cos(x)$ ha segno costante in tutto $\mathbb{R}^+$; ma la funzione coseno è definita anche dalla proprietà $\cos(0) = 1$, quindi si deve avere $\cos(x) > 0\ \forall x \in \mathbb{R}^+$.

Questo vuol dire che $\sin(x)$ è una funzione sempre crescente, poiché

$$\sin'(x) = \cos(x) > 0$$

In particolare $\sin(x) > 0\ \forall x \in \mathbb{R}^+$ perché $\sin(0) = 0$ e la funzione è crescente.

A sua volta questo implica che $\cos(x)$ è una funzione decrescente, poiché

$$\cos'(x) = -\sin(x) < 0$$

Dunque la funzione $\cos(x)$ è sempre positiva, vale $\cos(0) = 1$, ed è sempre decrescente. Allora deve avere un asintoto orizzontale (di valore compreso tra 0 e 1), quindi

$$\lim_{x\to+\infty} \cos'(x) = 0$$

D'altra parte $\cos'(x) = -\sin(x)$, quindi abbiamo

$$0 = \lim_{x\to+\infty} \cos'(x) = \lim_{x\to+\infty} -\sin(x)$$

$$\Longrightarrow \quad \lim_{x\to+\infty} \sin(x) = 0$$

il che è assurdo perché la funzione $\sin(x)$ è sempre crescente e vale $\sin(0) = 0$.

$\square$

Definizione 2.3.2 Avendo dimostrato che $\exists\, x > 0$ tale che $\cos(x) = 0$ definiamo il seguente simbolo:

$$\frac{\pi}{2} := \min\{x \geq 0 \mid \cos(x) = 0\}$$

Corollario $\sin\left(\frac{\pi}{2}\right) = 1$

Nota: la differenza tra questa definizione e quella data a lezione sta solo nel teorema che la precede, quindi al fine del programma svolto non cambia assolutamente niente. Ho voluto proporre quest'altro approccio perché la dimostrazione fatta così mi pare meno "arzigogolata"; a voi il giudizio.

2.3.2 Esponenziale e Logaritmo

11/03/2014

In maniera analoga a quanto fatto per le funzioni seno e coseno, possiamo definire la funzione esponenziale a partire da un Problema di Cauchy.

Dato il seguente problema di Cauchy:

$$\begin{cases} f' = f \\ f(0) = 1 \end{cases}$$

Definiamo la sua soluzione come $e^x := f(x)$, e studiamone le proprietà.

Proposizione 2.3.4 *Per ogni $x \in \mathbb{R}$ vale*

$$f(-x) = \frac{1}{f(x)}$$

Dimostrazione.

$$\frac{d}{dx}\left(f(x)f(-x)\right) = f'(x)f(-x) - f(x)f'(-x) = 0$$

Dunque la funzione $f(x)f(-x)$ è costante, con valore $f(x)f(-x) = f(0)f(-0) = 1 \quad \forall x$.

$\square$

Proposizione 2.3.5 *Per ogni $x, y \in \mathbb{R}$ vale*

$$f(x+y) = f(x)f(y)$$

Dimostrazione. Definiamo $g(x) = \frac{f(x+y)}{f(y)}$ e osserviamo che verifica lo stesso problema di Cauchy:

$$g'(x) = \frac{f(x+y)}{f(y)} = g(x)$$

$$g(0) = \frac{f(y)}{f(y)} = 1$$

Dunque $g(x) = f(x)$, che conclude la dimostrazione.

$\square$

Proposizione 2.3.6 *Per ogni $a \in \mathbb{R}$, e per ogni $n \in \mathbb{N}$ vale*

$$f(na) = (f(a))^n$$

Dimostrazione. Per induzione su n:

$\boxed{n = 1}$ ovvio

$\boxed{n \Rightarrow n+1}$

$$f((n+1)a) = f(na + a) = f(na)f(a) = (f(a))^n f(a) = (f(a))^{n+1}$$

$\square$

Proposizione 2.3.7 *f è strettamente crescente e valgono*

$$\lim_{x \longrightarrow +\infty} f(x) = \infty \qquad \lim_{x \longrightarrow -\infty} f(x) = 0$$

Dunque $f : \mathbb{R} \longrightarrow]0, +\infty[$ è bigettiva.

Dimostrazione. f è sempre positiva, poiché non è mai zero (possiamo scrivere $f(-x)$ come il reciproco), ma allora è crescente perché $f' = f$.

I due limiti vengono direttamente dalle proprietà dimostrate in precedenza.

□

Da tutte le proprietà dimostrate possiamo definire una funzione inversa della f:

$$\ln = f^{-1} :]0, +\infty[\longrightarrow \mathbb{R}$$

Che risulta bigettiva con derivata:

$$\ln'(y) = \left.\frac{1}{f'(x)}\right|_{x=\ln(y)} = \frac{1}{f(\ln(y))} = \frac{1}{y}$$

Proposizione 2.3.8 *Per ogni $x, y \in \mathbb{R}$ vale*

$$\ln(xy) = \ln(x) + \ln(y)$$

Dimostrazione. Definiamo $g(x) = \ln(xy) - \ln(x) - \ln(y)$ e calcoliamo la derivata:

$$g'(x) = \frac{1}{xy} y - \frac{1}{x} = 0$$

Dunque g è costante e vale $g(x) = g(1) = \ln(y) - \ln(1) - \ln(y) = 0 \ \ \forall x$.

□

2.4 Definizioni tramite integrali

2.4.1 Funzioni circolari

La funzione $g(t) = \frac{1}{\sqrt{1-t^2}}$ è integrabile in senso generalizzato in $[-1, 1]$.

Infatti per $t \longrightarrow 1^-$ risulta

$$g(t) = \frac{1}{\sqrt{1+t}} \frac{1}{\sqrt{1-t}} \simeq \frac{1}{2} \frac{1}{(1-t)^{\frac{1}{2}}}$$

e sappiamo che le funzioni del tipo $f(x) \simeq \frac{1}{(b-x)^\alpha}$ sono integrabili per $x \longrightarrow b^-$ se $\alpha < 1$ e non integrabili se $\alpha \geq 1$.

Visto che l'integrale converge possiamo dare un paio di definizioni.

Definizione 2.4.1 Definiamo innanzitutto la grandezza

$$\pi := 2\int_0^1 \frac{1}{\sqrt{1-t^2}}dt = \int_{-1}^1 \frac{1}{\sqrt{1-t^2}}dt$$

e per ogni $y \in [-1,1]$ definiamo la funzione

$$a(y) := \int_0^y \frac{1}{\sqrt{1-t^2}}dt$$

Studiamone le proprietà:
$a : [-1,1] \longrightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ è continua, bigettiva e dispari.
La sua derivata è $a'(y) = \frac{1}{\sqrt{1-y^2}} > 0 \ \forall y$, dunque la funzione è crescente.

Essendo bigettiva possiamo considerarne la funzione inversa:

$$\begin{array}{rcl} s : [-\frac{\pi}{2}, \frac{\pi}{2}] & \longrightarrow & [-1,1] \\ x & \longmapsto & a^{-1}(x) \end{array}$$

Usando quest'ultima funzione siamo in grado di definire, in modo diverso da quanto fatto fino ad ora, la funzione $\sin(x)$.

Definizione 2.4.2 Per $x \in \mathbb{R}$ definiamo

$$\sin(x) = \begin{cases} s(x) & \text{se } x \in [-\frac{\pi}{2}, \frac{\pi}{2}[\\ (-1)^k s(x-k\pi) & \text{se } x \in [k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}[\quad (k \in \mathbb{Z}) \end{cases}$$

Osservazione Per come è stata costruita la funzione risulta periodica di periodo 2π.

Studiamo la derivata della funzione $s(x)$: preso $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$

$$\frac{d}{dx}(s(x)) = \left.\frac{1}{\frac{d}{dy}a(y)}\right|_{y=s(x)} = \frac{1}{\frac{1}{\sqrt{1-s^2(x)}}} = \sqrt{1-s^2(x)}$$

$$\sin(x) = \begin{cases} s(x) & \text{se } x \in [-\frac{\pi}{2}, \frac{\pi}{2}[\\ s(\pi - x) & \text{se } x \in [\frac{\pi}{2}, \frac{3\pi}{2}[\end{cases}$$

Dunque la derivata della funzione seno risulta essere

$$\sin'(x) = \frac{d}{dx}(\sin(x)) = \begin{cases} \sqrt{1-s^2(x)} & \text{se } x \in]-\frac{\pi}{2}, \frac{\pi}{2}[\\ -\sqrt{1-s^2(x-\pi)} & \text{se } x \in]\frac{\pi}{2}, \frac{3\pi}{2}[\end{cases}$$

Da ciò che abbiamo trovato possiamo già trarre alcune conclusioni: la funzione $\sin(\cdot)$ è continua in $]-\frac{\pi}{2},\frac{3\pi}{2}[$, è derivabile in $]-\frac{\pi}{2},\frac{\pi}{2}[\cup]\frac{\pi}{2},\frac{3\pi}{2}[$ e vale $\lim_{x\longrightarrow\frac{\pi}{2}}\frac{d}{dx}(\sin(x))=0$

$$\implies \quad \frac{d}{dx}(\sin(x))\Big|_{x=\frac{\pi}{2}}=0$$

Lo studio della funzione derivata di $\sin(x)$ ci porta dunque alla seguente definizione:

Definizione 2.4.3 Per ogni $x\in\mathbb{R}$ definiamo la seguente funzione

$$\cos(x):=\frac{d}{dx}(\sin(x))=\sin'(x)$$

Osservazione Il coseno risulta essere una funzione pari periodica: $\cos(x+\pi)=-\cos(x)$, infatti $\sin(x+\pi)=-\sin(x)$ quindi

$$\cos(x+\pi)=\frac{d}{dx}(\sin(x+\pi))=\frac{d}{dx}(-\sin(x))=-\cos(x)$$

Studiamone la derivata analogamente a quanto fatto per la funzione seno:

Se $x\in[-\frac{\pi}{2},\frac{\pi}{2}]$ allora $\sin(x)=s(x)$, da cui $\cos(x)=\frac{d}{dx}(\sin(x))=\frac{d}{dx}(s(x))=\sqrt{1-s^2(x)}$. Allora possiamo calcolare

$$\frac{d}{dx}(\cos(x))=\frac{1}{2\sqrt{1-s^2(x)}}\big(-2\,s(x)\big)\sqrt{1-s^2(x)}=-s(x)=-\sin(x)$$

Se $x\notin[-\frac{\pi}{2},\frac{\pi}{2}]$ scriviamo $x=y+k\pi$, con $y\in[-\frac{\pi}{2},\frac{\pi}{2}[$. In questo caso

$$\frac{d}{dx}(\cos(x))=\frac{d}{dx}\big((-1)^k\cos(y)\big)=\frac{d}{dy}\big((-1)^k\cos(y)\big)\frac{dy}{dx}=\frac{d}{dy}\Big((-1)^k\cos(y)\Big)=$$
$$=(-1)^k(-s(y))=(-1)(-1)^ks(x-k\pi)=-\sin(x)$$

Proposizione 2.4.1 *Per ogni $x\in\mathbb{R}$ vale*

$$\sin^2(x)+\cos^2(x)=1$$

Dimostrazione. Definita la funzione $h(x):=\sin^2(x)+\cos^2(x)$,

essa risulta costante perché

$$h'(x)=2\sin(x)\cos(x)-2\cos(x)\sin(x)=0$$

Per trovare il valore basta quindi calcolarla in un punto:

$$h\left(\frac{\pi}{2}\right)=1+0=1$$

□

11/03/2014

Proposizione 2.4.2 (Formule di Addizione) *$\forall x \in \mathbb{R}$ valgono le seguenti relazioni:*

$$\begin{cases} \sin(x+y) = \sin(x)\cos(y) + \sin(y)\cos(x) \\ \cos(x+y) = \cos(x)\cos(y) - \sin(x) + \sin(y) \end{cases}$$

Dimostrazione. Dimostriamo solo la prima delle due uguaglianze, la seconda si fa in modo del tutto analogo.

Definiamo $F(x) = \sin(x+y) - \sin(x)\cos(y) - \cos(x)\sin(y)$ e mostriamo che è sempre nulla.

$$F'(x) = \cos(x+y) - \cos(x)\cos(y) + \sin(x)\sin(y)$$

$$F''(x) = -\sin(x+y) + \sin(x)\cos(y) + \cos(x)\sin(y) = -F(x)$$

$$\Longrightarrow F + F'' = 0$$

quindi $F^2 + F'^2$ è costante, perché:

$$(F^2 + F'^2)' = 2FF' + 2F'F'' = 2F'(F + F'') = 0$$

Calcoliamola nel punto $x = 0$: $F(0)^2 + F'(0)^2 = (\sin(y) - \sin(y))^2 + (\cos(y) - \cos(y))^2 = 0$.

Dunque la funzione $F^2 + F'^2$ è costantemente nulla, ma essendo la somma di due quadrati di funzioni reali questo implica che i due addendi sono identicamente nulli, quindi $F \equiv 0$, che completa la dimostrazione.

□

2.4.2 Esponenziale e Logaritmo

Definizione 2.4.4 Per $x > 0$ definiamo

$$\ln(x) = \int_1^x \frac{1}{t}\, dt$$

Il teorema fondamentale del calcolo integrale ci dice subito che

$$\ln'(x) = \frac{1}{x}$$

Studiamo le proprietà della nostra nuova funzione.

Proposizione 2.4.3 *Per ogni $x, y \in \mathbb{R}$ vale*

$$\ln(xy) = \ln(x) + \ln(y)$$

Dimostrazione. Usiamo la definizione di $\ln(\cdot)$:

$$\ln(xy) = \int_1^{xy} \frac{1}{t}\, dt = \int_1^{x} \frac{1}{t}\, dt + \int_x^{xy} \frac{1}{t}\, dt$$

Facendo un cambio di variabile, $s = \frac{t}{x}$, il secondo integrale diventa

$$\int_x^{xy} \frac{1}{t}\, dt \doteqdot \int_1^{y} \frac{1}{xs} x\, ds = \int_1^{y} \frac{1}{s}\, ds = \ln(y)$$

Quindi completiamo la dimostrazione:

$$\ln(xy) = \int_1^{x} \frac{1}{t}\, dt + \int_x^{xy} \frac{1}{t}\, dt = \ln(x) + \ln(y)$$

$\square$

Proposizione 2.4.4 *Per ogni $n \in \mathbb{Z}$ vale*

$$\ln(2^n) = n \ln(2)$$

e da ciò

$$\lim_{x \longrightarrow +\infty} \ln(x) = +\infty \ \textit{se } n > 0$$

$$\lim_{x \longrightarrow +0^+} \ln(x) = -\infty \ \textit{se } n < 0$$

Dimostrazione. Per induzione su n:
$\boxed{n = 1}$ $\ln(2) = \ln(2)$.
$\boxed{n \to n + 1}$

$$\ln(2^{n+1}) = \ln(2) + \ln(2^n) = \ln(2) + n \ln(2) = (n + 1) \ln(2)$$

I due limiti derivano direttamente dalla proprietà appena mostrata, perciò non li ridimostriamo.

$\square$

Una volta definito il logaritmo in questo modo possiamo definire la funzione esponenziale come sua inversa

$$\exp := \ln^{-1} : \mathbb{R} \longrightarrow]0, +\infty[$$

e ricostruire tutte le sue proprietà come abbiamo fatto in tutti gli altri casi.

2.5 Funzioni circolari alla Prodi

Definire seni e coseni alla maniera “classica”, ovvero come ascissa e ordinata di un punto nella circonferenza unitaria è un modo intuitivo che presenta nascosto il problema del calcolare la lunghezza di una curva, che è un problema tutt’altro che semplice!

Come si calcola la lunghezza di una curva?
Una *curva* è una funzione continua del tipo $\gamma : [a, b] \longrightarrow \mathbb{R}^2$.
Per calcolare la sua lunghezza prendiamo una scomposizione dell’intervallo $[a, b]$

$$\sigma = \{x_0, x_1, \ldots, x_n\} \subseteq [a, b]$$

$$a = x_0 < x_1 < \ldots < x_n = b$$

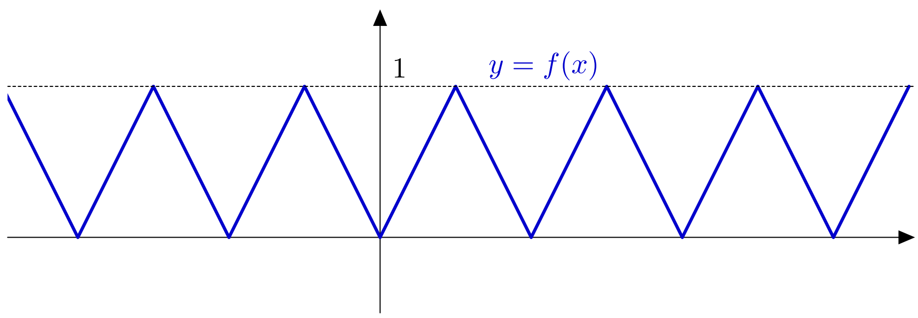
E consideriamo le figure poligonali inscritte in γ passanti per i punti dati. La lunghezza di una poligonale passante per quei punti è abbastanza semplice:

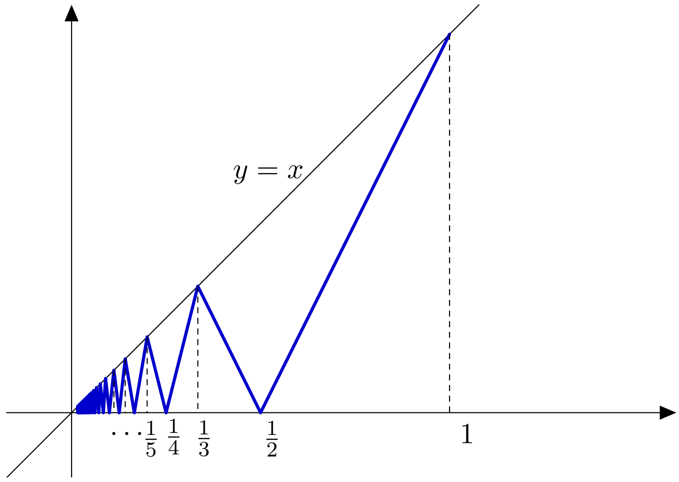
$$l(\mathcal{P}) = \sum_{k=1}^{n} |\gamma(x_k) - \gamma(x_{k-1})|$$

Allora definiamo la lunghezza della curva come

$$l(\gamma) := \sup\{l(\mathcal{P}) \mid \mathcal{P} = \text{ poligonale inscritta in } \gamma\}$$

In realtà richiedere solo che la funzione γ sia continua da luogo ad esempi di curve inaspettate che rendono la trattazione piuttosto complessa.

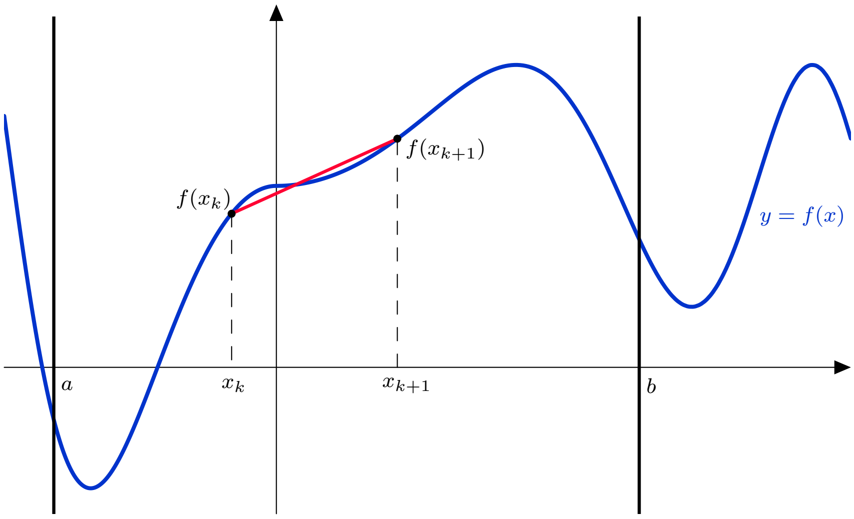




Se γ è C^1 allora possiamo prendere

$$l(\gamma) = \int_a^b \left|\gamma'(t)\right| dt$$

Per comprendere questa formula poniamoci nel caso in cui la nostra curva sia del tipo $\gamma = (x, f(x))$, cioè del tipo



In questo caso abbiamo che le lunghezze delle poligonali sono esprimibili come

$$l(\mathcal{P}) = \sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2} =$$

$$= \sum_{k=1}^{n} (x_k - x_{k-1}) \sqrt{1 + \left(\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}\right)^2} = \sum_{k=1}^{n} (x_k - x_{k-1}) \sqrt{1 + f'(t_k)^2}$$

In cui l'ultima uguaglianza è ottenuta tramite il teorema di Lagrange per un qualsiasi punto $t_k \in]x_{k-1}, x_k[$. A questo punto basta osservare che questa serie si può scrivere anche in forma integrale come

$$l(\mathcal{P}) = \sum_{k=1}^{n} (x_k - x_{k-1})\sqrt{1 + f'(t_k)^2} = \int_a^b \sqrt{1 + f'(t_k)^2}\, dx$$

2.6 Parametrizzare la circonferenza

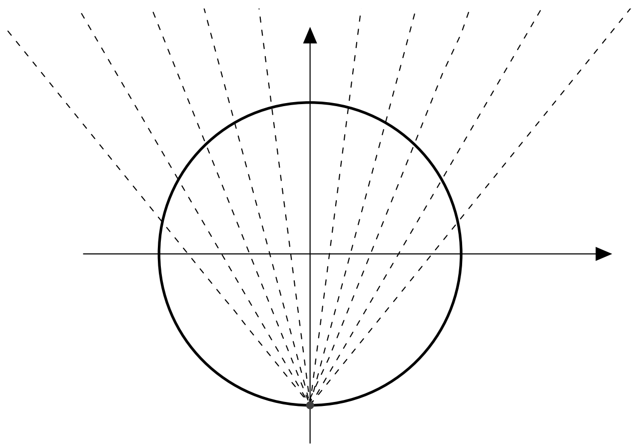
Abbiamo parlato nel paragrafo precedente delle curve e della loro lunghezza, o meglio della difficoltà di calcolarla. Uno dei metodi più efficaci per studiare una curva è quello di *parametrizzarla*, ma ancora una volta questo è un problema che potrebbe serbare qualche sorpresa.

Come possiamo parametrizzare una circonferenza?
La curva da cui partiamo è $\gamma = \{(x, y) \mid x^2 + y^2 = 1\}$.
La cosa più rapida sarebbe prendere

$$\begin{cases} x = \cos(t) \\ y = \sin(t) \end{cases}$$

Ma per costruire seni e coseni ci serve la circonferenza, quindi questo **non va bene**.

Proviamo un altro metodo, suggerito (o meglio schematizzato) dal seguente disegno:



Consideriamo la circonferenza unitaria γ, fissiamo il punto $x = (0, -1)$ e parametrizziamo la calotta superiore tramite le rette che congiungono i punti di essa col punto x scelto, ovvero intersechiamo le rette congiungenti con la circonferenza nel seguente modo:

$$\begin{cases} y = \frac{-1}{t}x - 1 \\ x^2 + y^2 = 1 \end{cases}$$

Troviamo i punti di intersezione:

$$x^2 + \frac{x^2}{t^2} + 1 + \frac{2x}{t} - 1 = 0 \quad \Longrightarrow \quad x\left(1 + \frac{1}{t^2}\right) + \frac{2}{t} = 0$$

$$\Longrightarrow \begin{cases} x = \frac{-2t}{1+t^2} \\ y = \frac{1-t^2}{1+t^2} \\ t \in [-1, 1] \end{cases}$$

E così è parametrizzata la calotta superiore.

In realtà se prendessimo $t \in \mathbb{R}$ potremmo parametrizzare tutta la circonferenza meno un punto, che è proprio il punto scelto da cui facciamo partire le rette.

Capitolo 3

Ricerca Operativa - Parte 1

17/03/2014

Iniziamo ora lo studio della **Ricerca Operativa**, anche detta *Teoria delle Decisioni*.

Le decisioni che possiamo prendere in generale possono essere in *condizioni di certezza*, ad esempio quando dobbiamo decidere se comprare un prodotto in un negozio a 10 euro o lo stesso prodotto in un altro negozio a 8 euro, in cui quindi la decisione conveniente è ovvia, o in *condizioni di incertezza*, di cui un esempio potrebbe essere decidere il voto da dare ad uno studente ad un esame orale, in cui i fattori da tenere in considerazione sono molti e non perfettamente delineati.

Sulle decisioni da prendere si ha quindi una **Relazione di Preferenza** descritta come $a \prec b$, che si legge "b è preferibile ad a", che rispetta le proprietà di antisimmetria e transitività.

Ci possono però essere decisioni non confrontabili, che chiameremo *indifferenti*.

Basandosi su queste possiamo definire una relazione d'equivalenza sullo spazio delle decisioni da prendere che chiameremo **Relazione di Indifferenza**.

Possiamo anche definire una funzione "valore"

$$w : \mathcal{C} \longrightarrow \mathbb{R}$$

che rispetta la proprietà

$$a \prec b \Longleftrightarrow w(a) < w(b)$$

in cui l'insieme $\mathcal{C}$ è l'insieme delle conseguenze.

3.1 Programmazione Lineare

Una parte molto importante della ricerca operativa è rappresentata dalla **Programmazione Lineare**, di cui parliamo in questo paragrafo.

Per capire di cosa si tratta supponiamo di essere un'azienda che produce determinati beni usando alcune materie prime. Ogni prodotto ha quindi un costo di produzione e un guadagno. Vorremmo massimizzare i guadagni netti, come possiamo fare?

Poniamoci all'inizio in una situazione molto semplice: un'azienda che produce due tipi di prodotti (ad esempio giocattoli di plastica) con due soli tipi di materie prime (plastica, ore di lavoro degli operai). Chiaramente le materie prime non le supponiamo infinite, ma con un determinato limite.

Definiamo quindi i dati su cui ci basiamo:

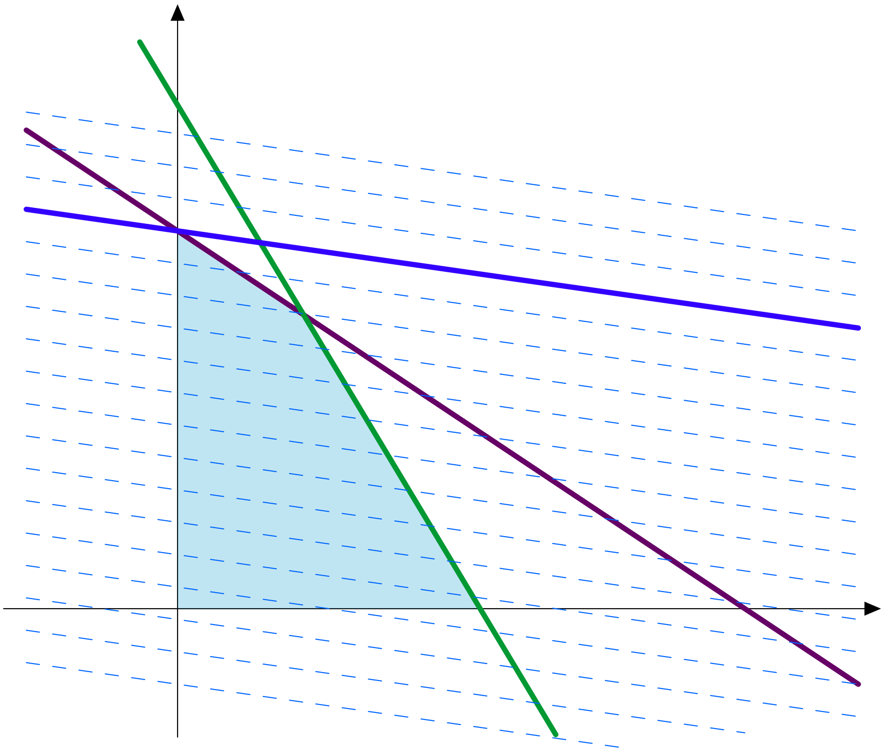
a_{11} = quantità di materia prima R_1 (plastica) necessaria a produrre un pezzo di tipo 1
a_{12} = quantità di materia prima R_1 (plastica) necessaria a produrre un pezzo di tipo 2
a_{21} = quantità di materia prima R_2 (ore) necessaria a produrre un pezzo di tipo 1
a_{22} = quantità di materia prima R_2 (ore) necessaria a produrre un pezzo di tipo 2
x_1 = numero di pezzi prodotti di tipo 1
x_2 = numero di pezzi prodotti di tipo 2
Funzione di Produzione: $p(x_1, x_2) = c_1x_1 + c_2x_2$
Limitazioni: $R_1 \leq r_1$, $R_2 \leq r_2$

Quindi dobbiamo cercare un risultato che verifichi il seguente sistema:

$$\begin{cases} \max\ p(x_1, x_2) \\ a_{11}x_1 + a_{12}x_2 \leq r_1 \\ a_{21}x_1 + a_{22}x_2 \leq r_2 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Le prime due equazioni a sistema (posto $\leq$ come $=$) rappresentano due rette nel piano cartesiano (x_1, x_2), quindi il sistema rappresenta una regione del piano che è la nostra zona di "scelte possibili".

Nello stesso piano (x_1, x_2) la funzione $p(x_1, x_2)$ rappresenta un fascio di rette parallele, per cui massimizzare la funzione vuol dire cercare quale di queste rette interseca la regione di piano individuata solo in un punto con coordinate massime.



Il nostro problema può quindi essere affrontato da un punto di vista geometrico sul piano (x_1, x_2). La stessa cosa si può fare se le grandezze in gioco sono tre, anche se il problema si presenterà in uno spazio tridimensionale piuttosto che in un piano.

Ma questo approccio non sarà possibile in generale quando le grandezze in gioco sono più di tre, quindi dobbiamo usare un diverso metodo.

Capitolo 4

Finanza Matematica

4.1 Leggi di Capitalizzazione

Ci si pone una scelta: accettare 100 euro oggi o 1000 euro tra un anno.

Chiaramente se la scelta fosse accettarne 100 oggi o 100 tra un anno non avremmo avuto dubbi, ma nel nostro caso non è così ovvio.

Per scegliere dobbiamo giustamente stabilire dei criteri, altrimenti non ne usciamo vivi. Iniziamo col definire le quantità che intervengono in questo caso.

C = *capitale* disponibile ora.
M = valore assunto al tempo t futuro (*montante*).
$M - C = I =$ *interesse*
$\frac{M}{C}$ = fattore di interesse
Legge di Capitalizzazione= $f(C,t)$ che dà M in funzione di C

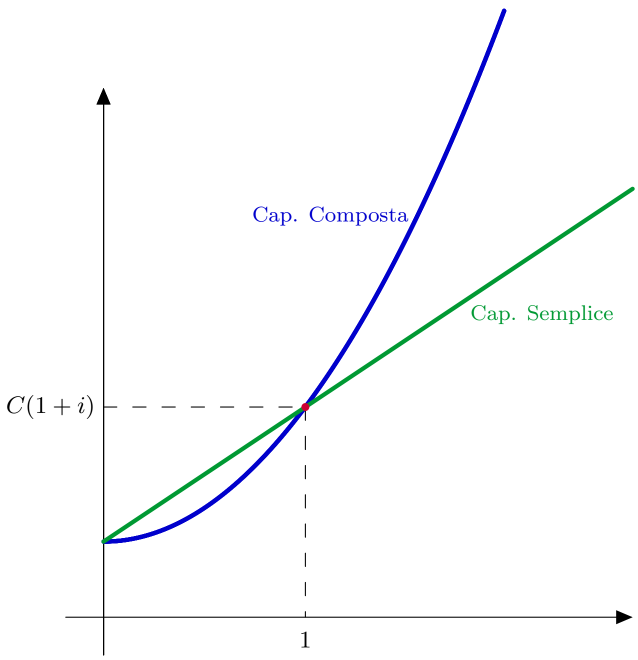
Le leggi di capitalizzazione possibili sono due:

1. Legge di Capitalizzazione *Semplice* con tasso annuo[1] i: $M = C(1 + it)$

2. Legge di Capitalizzazione *Composta*: $M = C(1 + i)^t$

Le due leggi si possono comparare graficamente:

[1] Qui si suppone di misurare il tempo t in anni, come viene solitamente fatto per ragioni di convenzione sociale, ma nulla ci vieta in teoria di usare un'altra misura.



Esempio 4.1.1 Un capitale di 1 euro viene investito con tasso $i = 10\% = 0.1$ in un tempo $t =$ (3 anni e 2 mesi)$=\frac{19}{6}$ anni. Con la legge di capitalizzazione composta M risulta essere

$$M = (1 + 0.1)^{\frac{19}{6}} = (1.1)^{\frac{19}{6}}$$

Prima delle calcolatrici il risultato era di difficile calcolo, e veniva quindi approssimato usando la capitalizzazione semplice nella parte frazionaria:

$$M = (1.1)^3 \cdot \left(1 + \frac{1}{6} \cdot 0.1\right)$$

Il calcolo fatto in questo modo viene chiamata *convenzione lineare*, che corrisponde ad approssimare la funzione esponenziale della legge di capitalizzazione composta con la funzione lineare tra i punti con ascissa intera.

Una volta stabilita la legge di capitalizzazione usata potremmo anche voler risolvere il problema inverso: avendo un importo attuale S disponibile al tempo t, quant'è il valore V tale che $f(V, t) = S$?

Questo dipende dalla legge di capitalizzazione scelta, quindi abbiamo due casi:

1. Capitalizzazione Composta:

$$V(1 + i)^t = S \Longrightarrow \quad V = S(1 + i)^{-t}$$

2. Capitalizzazione Semplice:

$$V(1+it) = S \Longrightarrow \quad V = S(1+it)^{-1}$$

La quantità $S - V$ si chiama *sconto*.

Nota: In termini numerici la quantità $S - V$ è esattamente la stessa della quantità $M - C$ usata sopra, quindi perché cambiargli nome? Il fatto è che si vuole sottolineare la differenza dei due problemi, che sono uno l'inverso dell'altro, ma quantitativamente non cambia niente.

Osservazione La legge di capitalizzazione composta è *Scindibile*, ovvero interrompendo l'impiego e riprendendolo immediatamente non si hanno variazioni del risultato.

Cosa vuol dire: supponiamo di aver investito 1 euro. Se decidiamo di ritirarli al tempo t_1 e di reinvestirli immediatamente, al tempo $t_2 > t_1$ è come se non li avessimo mai ritirati. Questo perché risulta

$$M_{t_2} = (1+i)^{t_2}$$
$$M_{t_1} M_{t_2-t_1} = (1+i)^{t_1}(1+i)^{t_2-t_1} = (1+i)^{t_2} = M_{t_2}$$

Questo non è vero per la legge di capitalizzazione semplice! Le quantità stavolta le possiamo calcolare come

$$M_{t_2} = (1+t_2 i)$$
$$\begin{aligned} M_{t_1} M_{t_2-t_1} &= (1+t_1 i)(1+(t_2-t_1)i) = \\ &= 1 + t_1 i + (t_2-t_1)i + t_1(t_2-t_1)i^2 = \\ &= 1 + t_2 i + t_1(t_2-t_1)i^2 = M_{t_2} + t_1(t_2-t_1)i^2 > M_{t_2} \end{aligned}$$

4.2 Tasso Nominale Annuo

In questo paragrafo parleremo di **Tasso Nominale Annuo convertibile k volte in un anno**, indicato con j_k.

In generale avere tasso nominale annuo j_k vuol dire avere un tasso $i_k = \frac{j_k}{k}$ per k volte durante un intero anno.

Ma nella pratica avere un interesse annuo fissato o un tasso nominale annuo cosa cambia?

Avere un tasso di $j_2 = 0.2$ vuol dire avere un tasso semestrale di $i_2 = \frac{0.2}{2}$.

Investire 1 euro con tasso del 20% annuo convertibile semestralmente, mi porta dopo un anno ad avere $(1 + i_2)^2 = 1.21$ euro, mentre con un interesse annuo del 20% avremmo avuto 1.20 euro.

Dunque si guadagna un po' di più che non un interesse annuo secco, ma quanto di più? Calcoliamocelo.

Un tasso nominale annuo j_k a quale tasso annuale (o interesse annuale) equivale?

$$\left(1 + \frac{j_k}{k}\right)^k = 1 + l \Longrightarrow l = \left(1 + \frac{j_k}{k}\right)^k - 1$$

Qual'è il tasso nominale annuo i_k equivalente al tasso annuo i?

$$(1 + i_k)^k = 1 + i \Longrightarrow i_k = \sqrt[k]{1 + i} - 1$$

Esempio 4.2.1 A quale tasso annuo equivale il 10% mensile?

Se dovessimo usare il semplice ragionamento ci verrebbe forse in mente di suggerire il 120%, visto che in fondo quello che si fa è di applicare 12 volte durante l'anno un tasso del 10% sul capitale investito.

In realtà questo ragionamento non è corretto, perché facendo i conti giusti l'interesse risulta

$$i_{12} = 0.1 \Longrightarrow \qquad i = (1.1)^{12} - 1 = 2.1384$$

Quindi il 10% mensile equivale ad un 213% annuo!

Come mai è così tanto più alto? Bé il fatto è che ogni volta che applichiamo il tasso nominale del 10% esso non viene applicato alla quantità di denaro investita inizialmente, ma alla quantità di denaro maturata fino a quel momento, quindi tende a crescere sempre di più.

18/03/2014

Quando si parla di tasso nominale annuo salta spesso fuori un diverso tipo di capitalizzazione rispetto a quelli visti in precedenza: la **capitalizzazione istantanea** di un tasso nominale annuo i. Per capire di cosa si tratta poniamoci in questo ambito:

La banca ci dice che il capitale investito ce lo capitalizza con tasso annuo i, ma divisibile durante l'anno quante volte vogliamo: 2,10,1000,...

Per capire cosa succede proviamo a vedere cosa succede quando decidiamo numeri diversi di divisione della capitalizzazione.

Se decidiamo di fare la conversione una sola volta all'anno il montante sarà $1 + i$.
Se decidiamo di farla due volte l'anno sarà $\left(1 + \frac{i}{2}\right)^2$.
In generale, se decidiamo di farla n volte l'anno sarà $\left(1 + \frac{i}{n}\right)^n$.
Portando al limite questa possibilità avremo

$$\lim_{n \longrightarrow +\infty} \left(1 + \frac{i}{n}\right)^n = e^i$$

Vediamo quant'è la differenza: il tasso annuo effettivo corrispondente a $i = 0.2$ convertibile 2 volte in un anno è di

$$M = (1.1)^2 = 1.21 \quad \Longrightarrow \quad I = 1.21 - 1 = 0.21 = i^*$$

Nel caso della capitalizzazione istantanea avremo

$$i^* = I = e^i - 1 = e^{0.2} - 1 = 0.2214$$

quindi il tasso passa dal 20% al 22%.

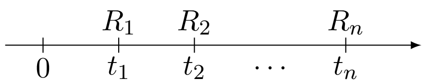
Se avessimo preso un tasso annuo del 10% saremmo passati al 10.5%.

Quindi l'incremento diventa sempre più alto all'aumentare del tasso nominale, come ci si aspetta da una funzione di tipo esponenziale come e^i.

4.3 Rendite

4.3.1 Rendite periodiche

Un'altra definizione utile è quella di **rendite**: una *rendita* è una sequenza di pagamenti disponibili in tempi futuri prefissati, che può essere rappresentata da un diagramma di questo tipo:



Se ad un certo momento decidiamo di ritirare del denaro, il valore attuale V dell'importo disponibile al tempo t è calcolabile come

$$V = \sum_{k=1}^{n} R_k (1 + i)^{-t_k}$$

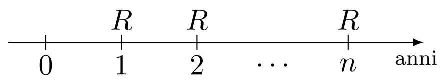
Anche il montante M è calcolabile, ma lo calcoliamo solo al termine di tutte le rate della rendita, perché l'idea è di calcolare di quanto disporremo al termine dei pagamenti prestabiliti, tenendo in considerazione tutti gli interessi fruttati da esse.

$$M = \sum_{k=1}^{n} R_k (1 + i)^{t_n - t_k} = V(1 + i)^{t_n}$$

Le rendite possono essere di due tipi:

1. Rendita *Periodica* = intervalli di tempo uguali
2. Rendita *Costante* = rate di uguale importo

Poniamoci nel caso di una rendita periodica e costante



Possiamo calcolare i valori che ci interessano

$$V = R\sum_{k=1}^{n}(1+i)^{-k} = R \cdot \frac{1-(i+1)^{-n}}{i} := R \cdot a_{n\rceil i}$$

in cui il valore $a_{n\rceil i}$ si chiama *a figurato n al tasso i*.

Il montante M dopo n anni sarà

$$M = V(1+i)^n = R \cdot \frac{(1+i)^n - 1}{i} = R \cdot s_{n\rceil i}$$

Esempio 4.3.1 Con i valori $V = 1000$, $R = 100$ e $n = 15$, come troviamo il tasso i?

$$1000 = 100\frac{1-(1+i)^{-15}}{i} = 100 \cdot a_{15\rceil i}$$

dunque $a_{15\rceil i} = 10$, ma per ricavare i da $a_{15\rceil i}$ a mano dovremo risolvere un'equazione di grado 15, il che non è proponibile.

Per fortuna la quantità $a_{15\rceil i}$ è molto usata in finanza, quindi possiamo ricavare il tasso i usando una tavola di valori (un computer sarebbe più preciso), che nel caso di $a_{15\rceil i} = 10$ restituisce un tasso $5 < i < 6$.

Esempio 4.3.2 (Tasse) Supponiamo di dover pagare un artigiano per un lavoro, e che lui ci proponga di decidere se pagare con o senza fattura. Facciamoci due conti:

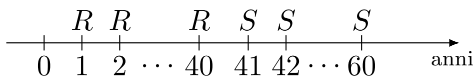
- Con Fattura: 20 euro, e la fattura renderà 1 euro all'anno allo stato per i successivi 10 anni.
- Senza Fattura: 20-V euro, in cui V è il valore della fattura.

Ma quant'è V?

$V = a_{\overline{10}|i}$, ma non sappiamo i. Se fosse $i = 3$ risulta $a_{\overline{10}|i} = 8.53$.

Quindi paradossalmente se pagassimo l'artigiano 20-8.53 euro e andassimo in banca dicendo di investire 8.53 euro in titoli di stato in modo che fruttino 1 euro ogni anno avremo fatto la stessa identica cosa.

Esempio 4.3.3 (La Pensione) Supponiamo di versare contributi per 40 anni con un tasso statale del 3% ($i = 0,03\%$). Dopo 40 anni il montante accumulato M mi verrà restituito in rate annue del valore S per 20 anni.



Calcoliamo i valori in gioco

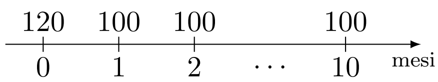
$$M = R \cdot s_{\overline{40}|0.03} \qquad\qquad V = S \cdot a_{\overline{20}|0.03}$$

$$\Longrightarrow \qquad S = R\frac{s_{\overline{40}|0.03}}{a_{\overline{20}|0.03}} = R\frac{75.40}{23.11} = 3.26$$

Ovvero, supponendo di pagare per 40 anni 10000 euro all'anno, avremmo diritto di ricevere 32600 euro all'anno per i successivi 20 anni.

La realtà non è così perché i titoli di stato non mantengono un tasso fisso per 40 anni, inoltre ogni individuo non paga solo i propri contributi, ma quelli di tutta la comunità.

Esempio 4.3.4 Supponiamo di comprare un televisore che costa 1000 euro in 10 rate mensili a tasso 0, ovvero in 10 rate mensili da 100 euro. Ma per aprire la pratica di tassazione mensile dobbiamo pagare 120 euro.



In questo caso abbiamo $V = 1000 = 120 + 880$, quindi $880 = 100 \cdot a_{\overline{10}|i}$ (ricordando che stavolta i è un tasso mensile, non annuale), da cui troviamo $a_{\overline{10}|i} = 8.8$ per ricavare il tasso mensile $i \simeq 0.03$.

Tutto sommato ci sembra abbastanza ragionevole, ma questo è un tasso mensile! Calcoliamoci il tasso annuale i^* derivante:

$$1 + i^* = (1.03)^{12} \quad \Longrightarrow \quad i^* = (1.03)^{12} - 1 = 0.42$$

Quindi il tasso annuo è del 42%, che non è affatto poco.

4.3.2 Rendite perpetue

21/03/2014

Ora parliamo di **rendite perpetue** di un importo R, ovvero una rendita costante ma infinita. In questo caso avremo dei valori diversi da quelli del caso precedente, calcolabili mandando al limite le somme finite:

$$V_n = R\sum_{k=1}^{n} \frac{1}{(1+i)^k} = R \cdot a_{n\rceil i}$$

$$V_\infty = R\sum_{k=1}^{\infty} \frac{1}{(1+i)^k} = R\frac{1}{1+i}\frac{1}{1-\frac{1}{1+i}} = \frac{R}{i}$$

Le rendite perpetue rappresentano una situazione di questo tipo: dopo aver vinto un ammontare V di euro, decidiamo di investirli in azioni e ogni anno riceviamo come rendita gli interessi maturati durante l'anno, che non reinvestiamo. A questo punto l'anno successivo l'ammontare di soldi investiti è sempre lo stesso, perché gli interessi li abbiamo ritirati, e dunque gli interessi saranno gli stessi[2], che anche stavolta ritiriamo come rendita.

Quindi la situazione in cui ci troviamo è quella di ricevere ogni anno la rendita Vi, costante nel tempo, per sempre.

4.4 Ammortamento di un prestito

Avendo parlato di rendite è il caso di parlare di un argomento che può essere considerato "duale" ad esso: l'ammortamento di un prestito di importo V in n rate annuali costanti R al tasso i.

Per far sì che il debito V si estingua nel tempo prestabilito e al tasso pattuito si dovrà avere che $V = R\,a_{n\rceil i}$, ovvero $R = \frac{V}{a_{n\rceil i}}$.

Esempio 4.4.1 Diciamo che Tizio ci presta $V = 100$ euro, da restituire in $n = 4$ rate al tasso $i = 0.05$. Questo vuol dire che per rimborsare il prestito dovremo dare rate che ammontano a

$$R = \frac{100}{a_{4\rceil 0.05}} = 28.20$$

Per estinguere un debito (prendiamo quello dell'esempio precedente) facciamo un **Piano di Ammortamento**:

Anni	*Q.I.*	*Q.C.*	*D.E.*	*D.R.*	*Rata*
0	0	0	0	100	0
1	5	23.20	23.20	76.80	28.20
2	3.84	24.36	47.56	52.44	28.20
3	2.62	25.58	73.14	26.86	28.20
4	1.34	26.86	100	0	28.20

[2] Chiaramente è una situazione fittizia in cui gli interessi rimangono fissi negli anni, ma come esempio ci accontentiamo.

In cui le quantità in gioco sono:

Q.I.=Quota Interesse=(Debito Residuo)·(Tasso di Interesse)
Q.C.=Quota Capitale=(Rata)-(Quota Interesse)
D.E.=Debito Estinto=(Quota di Capitale)+(Debito Estinto precedente)
D.R.=Debito Residuo=(Debito Restante precedente)-(Quota Capitale)

Quello che abbiamo fatto è di calcolare le rate annue in modo da estinguere in 4 anni il debito iniziale e tutti gli interessi maturati in questo periodo di tempo.

Avremmo potuto decidere di estinguere il debito non con rate costanti, ma con Quota Capitale costante. Usando l'esempio di prima quello che potevamo fare era decidere di pagare ogni anno 25 euro di debito, così in 4 anni avremmo sicuramente estinto il debito iniziale di 100 euro.

Ma le rate annuali dovranno tenere conto anche degli interessi che si maturano nel tempo. Quindi calcolando anche gli interessi in gioco, al tasso fisso del 5%, con questo metodo la situazione sarebbe stata di questo tipo:

Anni	*Q.I.*	*Q.C.*	*D.E.*	*D.R.*	*Rata*
0	0	0	0	100	0
1	5	25	25	75	30
2	3.75	25	50	50	28.75
3	2.5	25	75	25	27.5
4	1.25	25	100	0	26.25

4.5 Decisioni finanziarie

In questo paragrafo mostreremo come la matematica finanziaria incontra la teoria delle decisioni di cui abbiamo parlato nei capitoli precedenti.

4.5.1 Scelta tra due rendite a tasso noto

La prima situazione che analizziamo è quella di una prestazione non monetaria retribuita con rate annue.

Esempio: andiamo a tagliare il prato ad un vicino, che come retribuzione ci promette 100 euro tra un anno e altri 100 tra due anni (scelta 1). La moglie propone invece di darci 210 euro tra due anni (scelta 2). Ci viene allora offerto di decidere quale retribuzione scegliere.

Qual'è la scelta più conveniente?

Per capire come prendere la nostra decisione continuiamo un po' la storia: Avendo bisogno di soldi subito, e non fra qualche anno, decidiamo di andare in banca e dire "Tizio mi ha promesso 100 euro tra un anno e altri 100 dopo due anni, mentre la moglie me ne ha promessi 210 tra due anni. Vorrei dei soldi subito; Se cedo a voi queste cifre, quanto mi potete dare adesso?

Quello che fa la banca è calcolare i tassi di interesse in gioco nelle due situazioni:

$$V_1 = \frac{100}{1+i} + \frac{100}{(1+i)^2} \qquad\qquad V_2 = \frac{210}{(1+i)^2}$$

La situazione a noi preferibile sarà quindi quella che ci rende un valore attuale V più alto. Vediamo quindi quando $V_1 > V_2$, al variare del tasso i:

$$\frac{100}{1+i} + \frac{100}{(1+i)^2} > \frac{2100}{(1+i)^2} \iff 100(1+i) + 100 > 210 \iff$$

$$\iff 100 + 100i > 10 \iff i > 0.1$$

La situazione è quindi schematizzata dal fatto che più il tasso di mercato è alto e più sarà conveniente ricevere dei soldi subito piuttosto che tra qualche anno, quindi sarà preferibile V_1 a V_2.

Se ci pensiamo un attimo questa conclusione sembra ragionevole: se il tasso di mercato (ovvero l'interesse sui capitali investiti) in questo momento è alto è meglio ricevere i soldi subito per poterli investire prima che questo si abbassi; viceversa, se il tasso di mercato è basso possiamo permetterci un po' più di calma negli investimenti, che comunque in questo momento frutterebbero poco.

4.5.2 Scelta tra investimenti diversi

Per proporre la seconda situazione scegliamo come esempio una situazione reale, seppur semplificata: quella dei titoli di stato.

In banca al giorno d'oggi (marzo 2014) è presente un "BTP 2023 al 9%" con prezzo 147.11.

In generale i BTP sono titoli di stato in cui investiamo un capitale iniziale che ci verrà restituito all'anno indicato (il 2023 in questo caso) e che in questi anni produrrà un interesse con un tasso annuo esplicito (in questo caso al 9%) che verrà mantenuto fisso per tutto il periodo di investimento.

Il prezzo (147.11 per noi) significa che per investire 100 euro, così da ricevere 9 euro di interesse ogni anno fino al 2023, e ricevere poi nel 2023 109 euro, dobbiamo in realtà pagare inizialmente 147.11 euro.

Questo potrebbe sembrare assurdo a prima vista, ma il ragionamento che fa lo stato è questo: quando fu reso disponibile il BTP 2023 (inizio anni '90) il prezzo era chiaramente quello base, ovvero per investire 100 euro se ne pagavano 100. Il tasso del 9% fissato è molto alto, ma l'investitore deve far fronte al fatto che vedrà i suoi investimenti salire solo dopo molti anni. Andando avanti con gli anni un interesse annuo del 9% fisso diventa sempre più conveniente per gli investimenti a breve termine, quindi il prezzo deve salire.

Se ci facciamo due conti possiamo trovare che in questo caso si ha
Rendimento lordo = 3.28%
Rendimento netto = 2.37%

La situazione è riassumibile in questo schema:

-147.11	9	9		9	109	
0	1	2	$\cdots$	8	9	anni
(2014)	(2015)	(2016)		(2022)	(2023)	

Quindi quello che abbiamo è che il tasso di rendimento lordo del 3.28 è quel numero i che verifica

$$147.11 = \frac{9}{1+i} + \frac{9}{(1+i)^2} + \ldots + \frac{9}{(1+i)^8} + \frac{109}{(1+i)^9}$$

Il rendimento netto è minore perché in realtà poi lo stato trattiene una parte dei rendimenti come tasse, quindi quello che prenderemo è in realtà meno del 9% netto.

Esempio 4.5.1 Dati i due investimenti descritti dai seguenti diagrammi

1)

-1000	550	650		
0	1	2	$\cdots$	anni

2)

-1000	550		1500		
0	1	$\cdots$	10	$\cdots$	anni

Quale dei due è più conveniente?

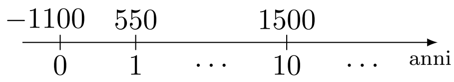
I tassi calcolati risultano essere $i_1 = 0.1268$, $i_2 = 0.1147$. Quindi se dovessimo giudicare la preferenza solo basandoci sui tassi il più conveniente risulta il primo.

Esempio 4.5.2 Se cambiamo gli investimenti iniziali le preferenze potrebbero cambiare di molto. Riprendiamo l'esempio precedente, cambiando gli investimenti iniziali:

1)

-1100	550	650		
0	1	2	$\cdots$	anni

2)



In questi casi i tassi risultano $i_1 = 0.0583$, $i_2 = 0.0963$. Quindi stavolta risulta preferibile il secondo.

Questo ci fa capire che investimenti a breve termine sono molto più influenzabili ai costi di investimento iniziale, che potevamo anche aspettarcelo.

Capitolo 5

Calcolo delle Variazioni

5.1 Nozioni principali

24/03/2014

Nella teoria del **Calcolo delle Variazioni**, che analizzeremo in questo capitolo, sono molto importanti alcuni risultati trovati dal famoso matematico **Pierre de Fermat**, quindi partiamo proprio da questo.

Teorema 5.1.1 (di Fermat (1)) *Preso un intervallo I, un punto $x_0 \in I$ estremante per una funzione $f : I \longrightarrow \mathbb{R}$, che risulta derivabile x_0, vale $f'(x_0) = 0$.*

Facciamo alcune osservazioni su questo risultato:

- Il fatto che $f'(x_0) = 0$ ci dice in particolare che

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \in \mathbb{R}$$

- La funzione f risulta differenziabile in x_0 se e solo se $\exists\, m \in \mathbb{R}$ tale che

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - m(x - x_0)}{x - x_0}$$

Dalla definizione di derivabilità proviene direttamente il concetto di "retta tangente" ad $f(x)$ nel punto x_0, mentre per il concetto di "piano tangente" a $z = f(x, y)$ nel punto $z_0 = (x_0, y_0, f(x_0, y_0))$ c'è da lavorare un po' di più.

Scriviamo $z = f(x_0, y_0) + m(x - x_0) + n(y - y_0)$ tali che

$$\lim_{(x,y) \to (x_0,y_0)} \frac{f(x, y) - \big(f(x_0, y_0) + m(x - x_0) + n(y - y_0)\big)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0$$

Inoltre f si dice *differenziabile* in (x_0, y_0) se $\exists\, L(h,k) = mh + nk$ lineare tale che

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - L(x-x_0, y-y_0)}{\| (x,y) - (x_0,y_0) \|} = 0$$

Osservazione f differenziabile in $(x_0, y_0) \Longrightarrow \exists\, f_x(x_0,y_0)$, $f_y(x_0,y_0)$ e $(m,n) = grad(f(x_0,y_0))$.

25/03/2014

Definizione 5.1.1 Dati X, Y spazi normati $f : X \longrightarrow Y$ si dice *differenziabile* in $x_0 \in X$ se $\exists\, L \in \mathcal{L}(X,Y)$ tale che

$$f(x) = f(x_0) + L(x - x_0) + o(\| x - x_0 \|)$$

$$\text{in cui vale} \qquad \lim_{\|x-x_0\|\longrightarrow 0} \frac{\| f(x) - f(x_0) - L(x-x_0) \|_Y}{\| x - x_0 \|_X}$$

L si dice *differenziale di f*, e si indica solitamente con "*df*".

Osservazione Se lo spazio di partenza è $X = \mathbb{R}^n$ le funzioni lineari $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ sono del tipo

$$L(h) = m_1h_1 + m_2h_2 + \ldots + m_nh_n$$

in cui h_i sono le componenti del vettore $h \in \mathbb{R}^n$.

La funzione lineare nulla $L = \underline{0}$ è del tipo $L(h) = 0\ \forall h \in \mathbb{R}^n$, quindi

$$L = \underline{0} \Longleftrightarrow m_1 = m_2 = \ldots = m_n = 0$$

Teorema 5.1.2 (di Fermat (2)) *Dato A un aperto di $\mathbb{R}$, $x_0 \in A$ e la funzione $f : A \longrightarrow \mathbb{R}$, se f è differenziabile in x_0, e x_0 è estremante*[1] *per f, allora il differenziale di f in x_0 è nullo: $df(x_0) = 0$.*

Dimostrazione. Poniamo che x_0 sia un punto di minimo per f.

Sia $L = df(x_0)$, e sia $y \in \mathbb{R} \setminus \{0\}$. Vogliamo provare che $L(y) = 0$.

Presa g la funzione reale definita $g(t) = f(x_0 + ty)$, per t "abbastanza piccolo" vale $g(t) \geq g(0)$ perché x_0 è un minimo per f.

Allora $g'(0) = 0$ perché 0 è un minimo relativo per g, dunque

$$\frac{g(t)-g(0)}{t} = \frac{1}{t}\Big(f(x_0+ty) - f(x_0)\Big) = \frac{1}{t}\Big(L(ty) + o(\| ty \|)\Big) =$$

$$= \frac{1}{t}\Big(tL(y) + o(t)\Big) = L(y) + o(1)$$

da cui otteniamo

$$0 = g'(0) = \lim_{t\longrightarrow 0} \frac{g(t)-g(0)}{t} = L(y)$$

□

[1] Ovvero punto di minimo o di massimo relativo.

Teorema 5.1.3 *Preso $X = \mathcal{C}^1([a,b])$, $f : [a,b] \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ continua, con derivate prime e seconde continue rispetto al secondo e terzo argomento*[2].

Allora la funzione $J : X \longrightarrow \mathbb{R}$ definita come

$$J(y) := \int_a^b f\big(x, y(x), y'(x)\big)\ dx$$

è differenziabile in ogni $y \in X$, ovvero $\exists\, L : x \longrightarrow \mathbb{R}$ lineare continua tale che

$$J(y+h) = J(y) + L(h) + o(\| h \|), \quad h \in X$$

Dimostrazione.

$$J(y+h) - J(y) = \int_a^b \Big(f\big(x, y+h, y'+h'\big) - f\big(x, y, y'\big) \Big)\ dx$$

Cerchiamo di esprimere la funzione sotto l'integrale in modo più comodo. Per farlo osserviamo la seguente cosa:

$$f(x, y+u, z+v) - f(x,y,z) = f'_y(x,y,z) \cdot u + f'_z(x,y,z) \cdot v + o\big(\| (u,v) \| \big)$$

quindi possiamo scrivere

$$J(y+h) - J(y) = \int_a^b \Big(f\big(x, y+h, y'+h'\big) - f\big(x, y, y'\big) \Big)\ dx =$$

$$= \int_a^b \Big(f'_y(x,y,y') \cdot h + f'_{y'}(x,y,y') \cdot h' + o(\| h \|) \Big)\ dx =$$

$$= \int_a^b \Big(f'_y(x,y,y') \cdot h + f'_{y'}(x,y,y') \cdot h' \Big)\ dx + o(\| h \|)$$

Il differenziale che abbiamo trovato ci dice inoltre che $dJ(y) = 0$ se $\forall h \in \mathcal{C}^1([a,b])$ vale

$$\int_a^b \Big(f'_y(x,y,y') \cdot h + f'_{y'}(x,y,y') \cdot h' \Big)\ dx = 0$$

per scrivere meglio questa equazione usiamo un paio di lemmi che ora enunciamo:

Lemma (1) *Sia $\alpha \in \mathcal{C}([a,b])$ tale che $\forall h \in \mathcal{C}_0([a,b])$* [3] *vale*

$$\int_a^b \alpha h = 0$$

Allora $\alpha(x) = 0\ \ \forall x \in [a,b]$.

---

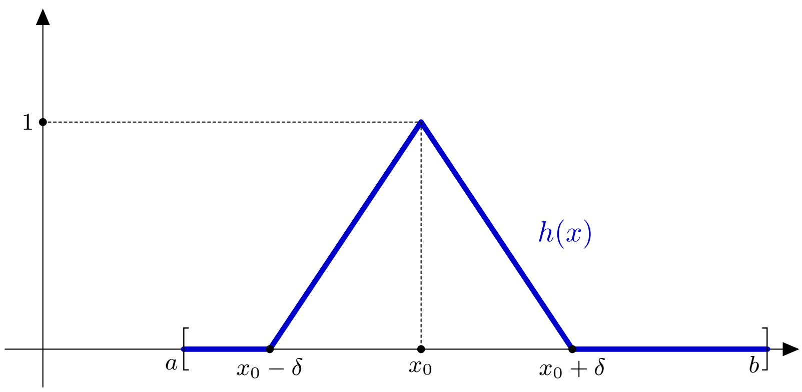
[2] Visto che f ha come dominio $[a,b] \times \mathbb{R}^2$, sarà del tipo $f(z,x,y)$.

[3] Con il simbolo $\mathcal{C}_0([a,b])$ indichiamo le funzioni continue nell'intervallo $[a,b]$ che valgono 0 agli estremi.

Dimostrazione. Se esistesse $x_0 \in [a,b]$ tale che $\alpha(x_0) > 0$ allora $\exists\, \delta > 0$ tale che

$$\alpha(x) > \frac{1}{2}\alpha(x_0) \quad \forall x \in [x_0 - \delta, x_0 + \delta]$$

Ma allora potremmo sempre trovare una funzione h tale per cui $\int_a^b \alpha \cdot h > 0$: basta ad esempio prendere una funzione h che sia positiva nell'intervallo $[x_0 - \delta, x_0 + \delta]$ ed ecco che il gioco è fatto.



Per la h in questo grafico vale infatti

$$\int_a^b \alpha h = \int_{x_0-\delta}^{x_0+\delta} \alpha h > \frac{1}{2}\alpha(x_0) \cdot \int_{x_0-\delta}^{x_0+\delta} h = \frac{1}{2}\alpha(x_0) \cdot \delta$$

Dunque l'ipotesi di partenza è assurda, che conclude la dimostrazione. □

Lemma (2) *Presa* $\alpha \in \mathcal{C}([a,b])$ *tale che* $\forall h \in \mathcal{C}_0^1([a,b])$ *vale*

$$\int_a^b \alpha h' = 0$$

Allora α *è costante in* $[a,b]$*.*

Dimostrazione. Scriviamoci la media integrale della nostra funzione

$$c = \frac{1}{b-a} \int_a^b \alpha(x)\, dx$$

Allora per definizione di c vale

$$\int_a^b \big(\alpha(x) - c\big)\, dx = 0$$

Chiamiamo poi

$$h(x) := \int_a^x \big(\alpha(t) - c\big)\, dt$$

che risulta derivabile, con derivata $h'(x) = \alpha(x) - c$, e tale che $h(a) = h(b) = 0$.

Allora $h \in \mathcal{C}_0^1([a,b])$, quindi per ipotesi $\int_a^b \alpha h' = 0$.

Proviamo a calcolare

$$\int_a^b \big(\alpha(x) - c\big)h'(x)\, dx = \int_a^b \alpha(x)h'(x)\, dx - c \cdot \int_a^b h'(x)\, dx =$$

$$= 0 - c \cdot \int_a^b h'(x)\, dx = -c\big(h(b) - h(a)\big) = 0$$

D'altra parte $h'(x) = \alpha(x) - c$, quindi abbiamo

$$0 = \int_a^b \big(\alpha(x) - c\big)h'(x)\, dx = \int_a^b \big(\alpha(x) - c\big)^2\, dx$$

Ma la funzione $\big(\alpha(x) - c\big)^2$ è positiva in tutto $[a,b]$, quindi se ha integrale nullo deve essere costantemente nulla, ovvero

$$\alpha(x) - c = 0 \ \ \forall x \qquad \Longrightarrow \qquad \alpha \equiv c$$

$\square$

Lemma (3) *Prese $\alpha, \beta \in \mathcal{C}([a,b])$ tali che $\forall h \in \mathcal{C}_0^1([a,b])$ vale*

$$\int_a^b (\alpha h + \beta h') = 0$$

Allora β è derivabile e vale $\beta' = \alpha$.

Dimostrazione. Sia

$$A(x) := \int_a^x \alpha(t)\, dt$$

cosicché $A'(x) = \alpha(x)$. Allora

$$\int_a^b \alpha h = \int_a^b A'h = \big[Ah\big]_a^b - \int_a^b Ah' = 0 - \int_a^b Ah'$$

in cui abbiamo usato l'integrazione per parti. Quindi

$$0 = \int_a^b (\alpha h + \beta h') = \int_a^b (-A + \beta)h'$$

Per il Lemma 2 questo implica che $-A + \beta$ è una funzione costante di valore c.

Allora $\beta = A + c$ è derivabile e risulta $\beta' = A' + 0 = \alpha$.

$\square$

Torniamo dunque al nostro problema: $L = dJ(y) = 0$ vuol dire che $\forall h \in \mathcal{C}^1[a,b]$ vale

$$\int_a^b \Big(f'_y(x,y,y') \cdot h + f'_{y'}(x,y,y') \cdot h' \Big)\, dx = 0$$

Per il Lemma 3, se chiamiamo $f'_y = \alpha$ e $f'_{y'} = \beta$, abbiamo che β è derivabile e risulta $\beta' = \alpha$, ovvero

$$\frac{d}{dx} f'_{y'}(x,y,y') = f'_y(x,y,y')$$

Questa equazione si chiama **Equazione di Eulero di** J.

$\square$

28/03/2014

Osservazione Se la funzione è del tipo $f(x,y')$, ovvero f non dipende da y, l'equazione di Eulero diventa

$$\frac{d}{dx} f'_{y'}(x,y') = k \in \mathbb{R}$$

Osservazione Se la funzione è del tipo $f(y,y')$, ovvero f non dipende da x, l'equazione di Eulero diventa

$$\frac{d}{dx} f'_{y'}(y,y') - f'_y(y,y') = 0$$

Possiamo calcolare anche le derivate seconde in certi argomenti, visto che avevamo supposto $f \in \mathcal{C}^2$ nel secondo e terzo argomento:

$$f''_{y'y}(y,y') \cdot y' + f''_{y'y'}(y,y') \cdot y'' - f'_y(y,y') = 0$$

$$\underbrace{f''_{y'y} \cdot y'^2 + f''_{y'y'} \cdot y''y' - f'_y \cdot y'}_{\frac{d}{dx}\left(f - y' \cdot f'_{y'}\right)} = 0$$

$$f - y' \cdot f'_{y'} = k \in \mathbb{R}$$

Esempio 5.1.1 Un esempio di funzione, o *funzionale*, J del teorema di sopra è la funzione di lunghezza del grafico di $y \in \mathcal{C}^1([a,b])$, definito come

$$J(y) = \int_a^b \sqrt{1 + (y'(x))^2}\, dx$$

in cui la funzione integranda è $f(x,y,y') = \sqrt{1 + y'^2}$.

In questo caso abbiamo

$$f'_y = 0 \qquad\qquad f'_{y'} = \frac{y'}{\sqrt{1 + y'^2}}$$

quindi l'equazione di Eulero ci dice che

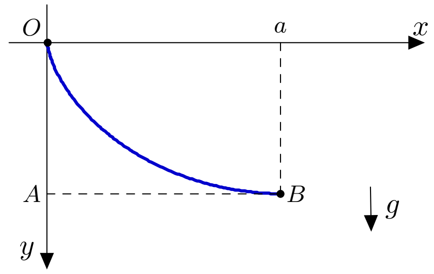
$$\frac{d}{dx} \frac{y'}{\sqrt{1+y'^2}} = 0 \quad \Longrightarrow \quad \frac{y'}{\sqrt{1+y'^2}} \equiv c$$

Il quale implica che y' è una funzione costante, quindi $y(x) = mx + q$ è una retta.

Questo ci dice che la curva che minimizza il funzionale di lunghezza è proprio la retta, ovvero che la curva di lunghezza minima che collega due punti è proprio una retta.

5.2 Cicloide o Brachistocrona

Presi due punti O, B nello spazio, vogliamo calcolare la curva che minimizza il tempo di percorrenza da O a B di un corpo in moto accelerato sotto l'effetto della forza gravità, ovvero cerchiamo la curva detta **Brachistocrona** da O a B.



Visto che ad ogni momento deve valere la legge di conservazione dell'energia dobbiamo avere

$$\frac{1}{2}mv^2 = mgy$$

In particolare il punto materiale si deve muovere secondo una certa legge oraria dipendente dal tempo

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases}$$

Calcoliamo la velocità

$$v = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dx}\right)^2 \left(\frac{dx}{dt}\right)^2} = \frac{dx}{dt}\sqrt{1 + y'^2}$$

quindi la nostra equazione diventa

$$\frac{1}{2}\left(\frac{dx}{dt}\right)^2 (1 + y'^2) = gy$$

$$\frac{dx}{dt} = \sqrt{\frac{2gy}{1 + y'^2}} \quad \Longrightarrow \quad \frac{dt}{dx} = \sqrt{\frac{1 + y'^2}{2gy}}$$

Abbiamo finalmente trovato l'espressione del tempo in funzione delle coordinate, quindi possiamo avere il tempo di percorrenza calcolando il tempo in funzione dell'ascissa del punto B, ovvero calcolare quando il punto materiale raggiunge B:

$$t(a) = \int_0^a \frac{dt}{dx} dx = \int_0^a \sqrt{\frac{1+y'^2}{y}} \sqrt{\frac{1}{2g}} dx$$

Quindi la funzione f che cerchiamo la possiamo esprimere come

$$J(y) = \int_0^a \sqrt{\frac{1+y'^2}{y}} dx \quad \Longrightarrow \quad f(x,y,y') = \sqrt{\frac{1+y'^2}{y}}$$

Possiamo calcolare l'equazione di Eulero ora che abbiamo la funzione:

$$f'_{y'} = \frac{1}{\sqrt{y}} \cdot \frac{2y'}{2\sqrt{1+y'^2}}$$

$$f - y' f'_{y'} = \frac{\sqrt{1+y'^2}}{\sqrt{y}} - \frac{y'^2}{\sqrt{y}\sqrt{1+y'^2}} = \frac{1+y'^2-y'^2}{\sqrt{y}\sqrt{1+y'^2}} = \frac{1}{\sqrt{y(1+y'^2)}}$$

Tutto questo deve essere eguagliata ad una costante, ma affinché l'espressione appena trovata sia costante basta che sia costante il denominatore, quindi per trovare la curva dovremo risolvere il seguente sistema:

$$\begin{cases} y(1+y'^2) = 2c \\ y(0) = 0 \\ y(a) = A \end{cases}$$

Per risolvere questa equazione differenziale possiamo porre $p(y') := \frac{2c}{1+y'^2}$, così da trasformare l'equazione $y(1+y'^2) = 2c$ semplicemente in $y = p(y')$.

Da ciò possiamo considerare la y' un parametro con cui trovare le equazioni parametriche della curva, ovvero

$$t = y' = \frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = p'(t)\frac{dt}{dx}$$
$$\Longrightarrow \quad \frac{dx}{dt} = \frac{1}{t}p'(t) = \frac{1}{t}\frac{-4c}{(1+t^2)^2}$$

Quindi possiamo trovare la parametrizzazione della coordinata x semplicemente tramite un integrale:

$$x(t) = \int \frac{-4c}{(1+t^2)^2} dt$$

Che si calcola tramite la sostituzione

$$t = \cot\left(\frac{\alpha}{2}\right) \qquad dt = \frac{-1}{2}\left(1 + \cot^2\left(\frac{\alpha}{2}\right)\right)$$

$$\implies \quad x(t) = \int \left(-4c \frac{1}{\left(1+\cot^2\left(\frac{\alpha}{2}\right)\right)^2} \cdot \left(\frac{-1}{2}\right) \left(1+\cot^2\left(\frac{\alpha}{2}\right)\right) \right) d\alpha =$$

$$= \int \frac{2c}{\frac{\sin^2(\frac{\alpha}{2})+\cos^2(\frac{\alpha}{2})}{\sin^2(\frac{\alpha}{2})}} \, d\alpha \;=\; c\int 2\sin^2\left(\frac{\alpha}{2}\right) \, d\alpha$$

Da cui otteniamo (eseguendo altri brevi calcoli) la seguente parametrizzazione:

$$\begin{cases} x(\alpha) = c(\alpha - \sin(\alpha)) + k \\ y(\alpha) = c(1-\cos(\alpha)) \end{cases}$$

Se calcoliamo i valori in 0 possiamo trovare il valore di k:

$$\begin{cases} x(0) = k \\ y(0) = 0 \end{cases} \implies k = 0$$

Ovvero

$$\frac{dy}{dx} = \frac{y'(\alpha)}{x'(\alpha)} = \frac{c\cdot\sin\alpha}{c(1-\cos\alpha)}$$

Quindi la curva che abbiamo trovato è la **Cicloide**, ovvero la curva generata da un punto fissato su una ruota che rotola con moto rettilineo uniforme di velocità v in una superficie orizzontale.

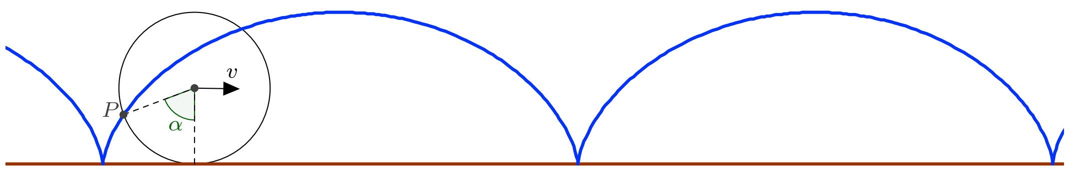


Figura 5.1: Cicloide

Possiamo infatti calcolare le coordinate del punto P che genera la curva:

$$\begin{cases} x(\alpha) = c(\alpha - \sin(\alpha)) \\ y(\alpha) = c(1-\cos(\alpha)) \end{cases}$$

e verificare che sono proprio quelle che avevamo trovato prima.

Quindi la curva che minimizza il tempo di percorrenza tra due punti (in assenza di forze esterne) è proprio la Cicloide. Proviamo a studiarne alcune proprietà:

31/03/2014

Fissato un arco di Cicloide, ad esempio quello con $\alpha \in [0, 2\pi]$, calcoliamone la **lunghezza**:

$$\int_0^{2\pi} \sqrt{(x'(\alpha))^2 + (y'(\alpha))^2}\, d\alpha = \int_0^{2\pi} \sqrt{1 + \cos^2\alpha - 2\cos\alpha + \sin^2\alpha}\, d\alpha =$$

$$= \int_0^{2\pi} \sqrt{2 - 2\cos\alpha}\, d\alpha = \int_0^{2\pi} 2\left|\sin\frac{\alpha}{2}\right|\, d\alpha = \left[-4\cos\frac{\alpha}{2}\right]_0^{2\pi} = 8$$

Quindi viene un numero intero, che è piuttosto singolare per una curva generata da funzioni circolari!

Per calcolare l'**area sottesa** ci sono invece diversi metodi equivalenti:

1. Se chiamiamo $g(\alpha) = 1 - \cos\alpha$, in linea di principio potremmo ricavare α dalla prima equazione $x = \alpha - \sin\alpha = p(\alpha)$ come $\alpha = p^{-1}(x)$, visto che $p(\cdot)$ risulta invertibile.

   Allora l'area è calcolabile tramite l'integrale

$$\int_0^{2\pi} g(\alpha(x))\, dx$$

   che possiamo calcolare tramite sostituzione:

$$\alpha(x) = t \qquad\qquad x = p(t) = t - \sin t$$

$$\Longrightarrow \int_0^{2\pi} g(\alpha(x))\, dx = \int_0^{2\pi} g(t)p'(t)\, dt = \int_0^{2\pi} (1 - \cos t)^2\, dt =$$

$$= \int_0^{2\pi} \big(1 + \cos^2 t - 2\cos t\big)\, dt = \int_0^{2\pi} \left(1 + \frac{1}{2}\big(1 + \cos(2t)\big) - 2\cos t\right)\, dt =$$

$$= \left[\frac{3}{2}t + \frac{1}{4}\sin(2t) - 2\sin t\right]_0^{2\pi} = 3\pi$$

2. Possiamo anche usare una teoria un po' più avanzata come quella degli integrali multipli, proviamo a farlo:

   Prendiamo l'area sottesa dalla cicloide, ovvero

$$A = \{(x, y) \in \mathbb{R}^2 \mid \exists\, \alpha \in [0, 2\pi],\ x = \alpha - \sin\alpha,\ 0 \leq y \leq 1 - \cos\alpha\}$$

   E usiamo il cambio di coordinate:

$$\begin{cases} x = \alpha - \sin\alpha \\ y = t \end{cases}$$

   trasformando l'insieme A nell'insieme B così definito:

$$B = \{(\alpha, t) \mid \alpha \in [0, 2\pi],\ 0 \leq t \leq 1 - cos\alpha\}$$

Allora la matrice Jacobiana del cambio di variabili risulta essere

$$J(\alpha, t) = \begin{pmatrix} 1-\cos\alpha & 0 \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad |det(J(\alpha, t)| = 1 - \cos\alpha$$

Quindi ora possiamo calcolare l'area di A semplicemente come

$$\iint_A 1 \; dxdy \doteq \iint_B (1-\cos\alpha) \; d\alpha dt =$$

$$= \int_0^{2\pi} \left(\int_0^{1-\cos\alpha} \big(1-\cos\alpha\big) \; dt \right) \; d\alpha = \int_0^{2\pi} \big(1-\cos\alpha\big)^2 \; d\alpha$$

che ci riporta ai calcoli precedenti, che non ripetiamo.

3. Esiste ancora un altro metodo per calcolare l'area della cicloide: se usiamo gli integrali curvilinei possiamo calcolare l'area dell'insieme A, definito prima, come

$$\int_{\partial A} -ydx = \int_0^{2\pi} \underbrace{(1-\cos\alpha)}_{y(\alpha)} \underbrace{(1-\cos\alpha)}_{x'(\alpha)} \; d\alpha$$

in cui il simbolo "∂A" indica il bordo dell'insieme A.

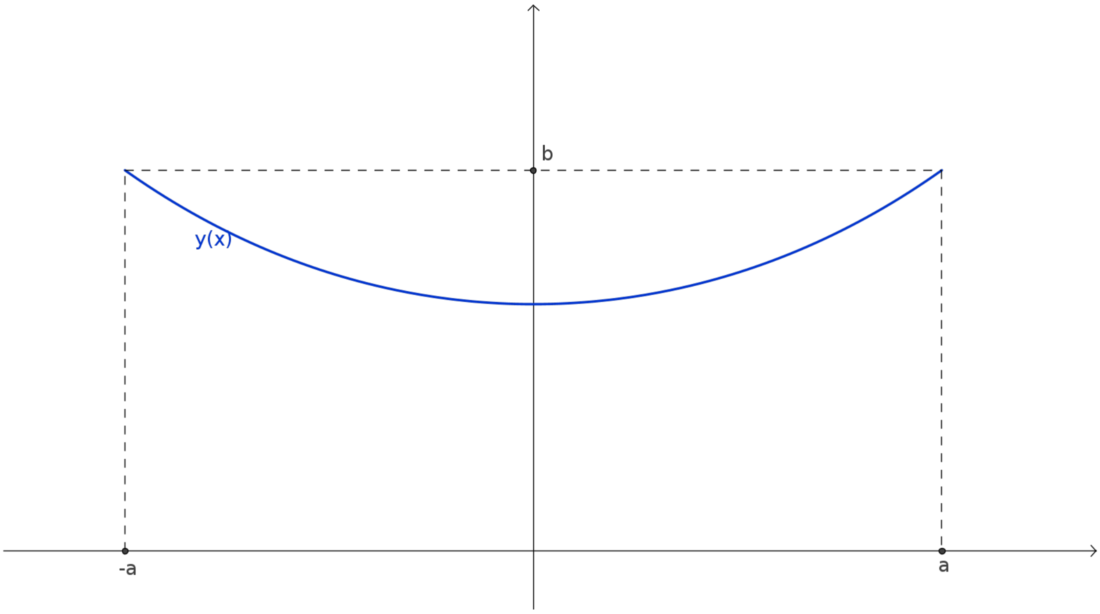
Ancora una volta questo metodo ci riporta agli stessi calcoli.

5.3 Catenaria

Supponiamo di fissare a due punti alla stessa altezza[4] un filo di massa non trascurabile, flessibile, inestensibile, e omogeneo. Che curva forma il filo?

Il problema fu proposto da Galileo nel 1638, e fu risolto dai fratelli Bernoulli solo due secoli dopo.

[4] Il fatto che i due punti siano presi alla stessa altezza semplifica un po' l'espressione finale della curva che otterremo, ma nulla ci vieta di prendere due punti qualsiasi e svolgere tutti gli studi allo stesso modo.



La lunghezza del filo si suppone fissata di valore $L > 2a$, visto che il filo è inestensibile. Inoltre la massa del filo è proporzionale alla sua lunghezza perché il filo è omogeneo.

Dobbiamo quindi minimizzare la funzione

$$J(y) := \int_{-a}^{a} y\sqrt{1+y'^2}\,dx$$

tra tutte le funzioni $y \in \mathcal{C}^\infty([-a,a])$ tali che $y(\pm a) = b$ e

$$\int_{-a}^{a} \sqrt{1+y'^2}\,dx = L$$

Questo problema ci ricorda i **Moltiplicatori di Lagrange**:

Proposizione 5.3.1 (Regola per i Moltiplicatori) *Prese due funzioni $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}$, con $f, g \in \mathcal{C}^1$.*

Se chiamiamo $A = \{x \in \mathbb{R}^n \mid g(x) = 0\}$, tale che grad$(g) \neq 0\ \forall x \in A$, e $\overline{x}$ estremante per $f\big|_A$, allora $\exists\, \lambda \in \mathbb{R}$ tale che grad$(f - \lambda g)(\overline{x}) = \underline{0}$.

Corollario (Estensione della Regola) *Prese due funzioni $f, g : [a,b] \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$, con le ipotesi dette prima per le funzioni $y \in \mathcal{C}^\infty$.*

Se chiamiamo $A = \{y \in \mathcal{C}^1([a,b]) \mid y(a) = a',\ y(b) = b', G(y) = C\}$, dove

$$G(y) := \int_a^b g(x, y, y')\,dx$$

tale che $dG \neq 0 \; \forall x \in A$, *e* $\overline{y}$ *estremante per* $J\big|_A$, *dove*

$$J(y) = \int_a^b f(x, y, y')\, dx$$

allora $\exists\, \lambda \in \mathbb{R}$ *tale che* $d(J + \lambda G) = 0$.

Quindi abbiamo

$$(J + \lambda G)(y) = \int_a^b \big(f(x, y, y') + \lambda g(x, y, y')\big)\, dx$$

Se chiamiamo $H(x, y, y') := f(x, y, y') + \lambda g(x, y, y')$ abbiamo che se $(J + \lambda G)(y) = 0$ allora deve essere soddisfatta L'Equazione di Eulero, ovvero deve valere

$$\frac{d}{dx} H'_{y'} = H'_y$$

D'altra parte, se H non dipende da x l'equazione diventa del tipo

$$H(y, y') - y' H'_{y'} = c \quad \text{(costante)}$$

Nel nostro caso

$$H(y, y') = y\sqrt{1 + y'^2} + \lambda\sqrt{1 + y'^2} = (y + \lambda)\sqrt{1 + y'^2}$$

quindi calcoliamo le derivate che ci servono:

$$H'_{y'} = (y + \lambda)\, \frac{y'}{\sqrt{1 + y'^2}}$$

che sostituiamo nell'equazione di Eulero per ottenere

$$H - y' H'_{y'} = (y + \lambda)\sqrt{1 + y'^2} - (y + \lambda)\, \frac{y'^2}{\sqrt{1 + y'^2}} =$$

$$= (y + \lambda)\, \frac{1 + y'^2 - y'^2}{\sqrt{1 + y'^2}} = \frac{y + \lambda}{\sqrt{1 + y'^2}}$$

e tutto quello che dobbiamo fare è porre questa espressione uguale ad una costante c:

$$\frac{y + \lambda}{\sqrt{1 + y'^2}} = c \quad \Longrightarrow \quad y = -\lambda + c\sqrt{1 + y'^2}$$

Se chiamiamo $p(y') := -\lambda + c\sqrt{1+y'^2}$ la nostra espressione è del tipo $y = p(y')$, allora possiamo considerare y' come un parametro t, proprio come abbiamo fatto nell'esempio della Brachistocrona, ed ottenere un problema del tipo

$$\begin{cases} x = x(t) \\ y = p(t) \end{cases}$$

di cui possiamo calcolare direttamente l'espressione di $x(t)$:

$$x'(t) = \frac{dx}{dt} = \frac{dx}{dy}\frac{dy}{dt} = \frac{1}{t}p'(t) = \frac{1}{t}c\frac{t}{\sqrt{1+t^2}} = \frac{c}{\sqrt{1+t^2}}$$

$$\Longrightarrow \quad x(t) = \int \frac{c}{\sqrt{1+t^2}}\,dt + k = c \cdot arcsinh(t) + k$$

Quindi il sistema è in realtà

$$\begin{cases} x(t) = c \cdot arcsinh(t) + k \\ y(t) = -\lambda + c\sqrt{1+t^2} \end{cases}$$

di cui possiamo esplicitare il parametro t:

$$arcsinh(t) = \frac{x-k}{c} \qquad \Longrightarrow \qquad t = \sinh\left(\frac{x-k}{c}\right)$$

$$y = -\lambda + \sqrt{1+t^2} = -\lambda + c\cosh\left(\frac{x-k}{c}\right)$$

01/04/2014

Adesso quello che dobbiamo fare è esplicitare i parametri c e k lasciati in sospeso.
Uno dei due è molto semplice da trovare: $y(x) = y(-x) \Longrightarrow k = 0$.
Per l'altro c'è da fare un po' più di fatica:

$$y = -\lambda + c\cosh\left(\frac{x}{c}\right) \Longrightarrow y' = \sinh\left(\frac{x}{c}\right)$$

Ricordiamo che la lunghezza della curva è esprimibile come

$$L = l(\gamma) = \int_{-a}^{a} \sqrt{1+\sinh^2\left(\frac{x}{c}\right)}\,dx = \int_{-a}^{a} \cosh\left(\frac{x}{c}\right)\,dx = 2c\sinh\left(\frac{a}{c}\right)$$

Dalla relazione $L = 2c\sinh\left(\frac{a}{c}\right)$ appena trovata, equivalente a $2\sinh\left(\frac{a}{c}\right) - \frac{L}{c} = 0$, vorremmo esplicitare c, che sappiamo essere positivo.

Definiamo $q(c) := 2\sinh\left(\frac{a}{c}\right) - \frac{L}{c}$ e studiamo questa funzione per trovare quando si annulla:

$$\lim_{c\longrightarrow+\infty} q(c) = 0 \qquad\qquad \lim_{c\longrightarrow0^+} q(c) = +\infty$$

$$q'(c) = 2\left(\frac{-a}{c^2}\right)\cosh\left(\frac{a}{c}\right) + \frac{L}{c^2} = \frac{1}{c^2}\left(L - 2a\cosh\left(\frac{a}{c}\right)\right)$$

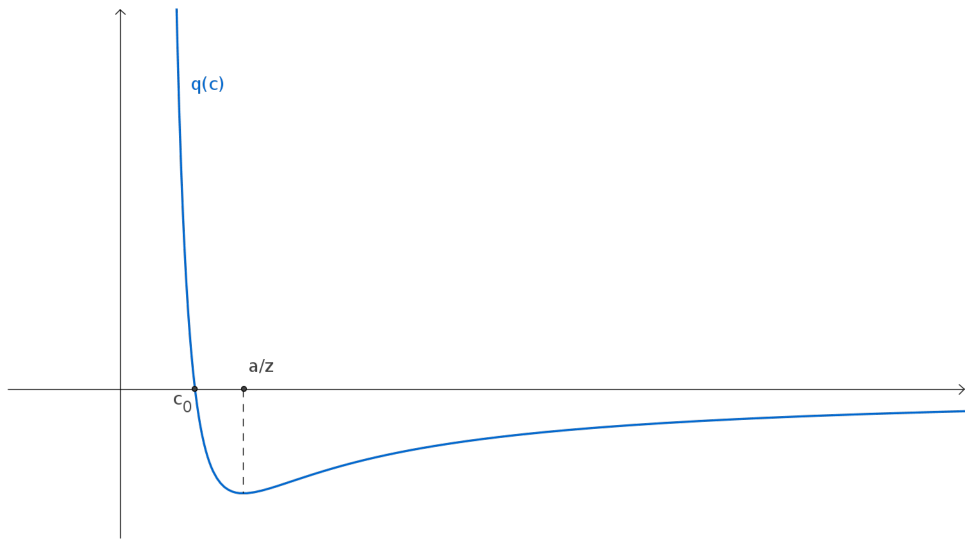
quindi $q'(c) = 0$ se $\cosh\left(\frac{a}{c}\right) = \frac{L}{2a}$, che è una condizione che avevamo già trovato in precedenza, ovvero

$$q'(c) = 0 \quad \Longleftrightarrow \quad \cosh\left(\frac{a}{c}\right) = \frac{L}{2a} \quad \Longleftrightarrow \quad \frac{a}{c} = z := arccosh\left(\frac{L}{2a}\right)$$

Mentre per la positività vale

$$q'(c) > 0 \quad \Longleftrightarrow \quad \frac{a}{c} < z \quad \Longleftrightarrow \quad c > \frac{a}{z}$$

Quindi la funzione $q(c)$ è di questo tipo:



5.4 Altri esempi

Facciamo ora un altro esempio di calcolo delle variazioni, meno classico rispetto ai primi due in cui si sono trovate la Cicloide e la Catenaria.

Esempio 5.4.1 Vogliamo minimizzare

$$J(y) = \int_{-1}^{1} \left(1 - y'(x)^2\right)^2 \, dx$$

con le condizioni $y(1) = y(-1) = 0$.

In questo caso la funzione f è data da $f(x, y, y') = \big(1 - y'(x)^2\big)^2$, quindi l'equazione di Eulero è del tipo

$$\frac{d}{dx} f'_{y'} = f'_y$$

Ma $f'_y = 0$ perché la funzione f dipende solo da y', dunque l'equazione di Eulero ci dice semplicemente che $f'_{y'}$ deve essere costante.

$$c = f'_{y'} = 2(1 - y'^2) \cdot 2y' \quad \Longleftrightarrow \quad (1 - y'^2)y' = c$$

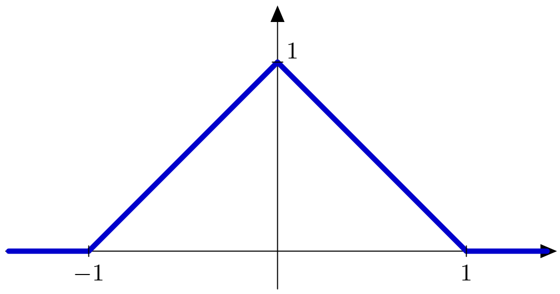
Ma l'espressione $(1 - y'^2)y'$ è costante se e solo se y' è costante, quindi dobbiamo avere $y' = m$, ovvero

$$\begin{cases} y(x) = mx + q \\ y(1) = y(-1) = 0 \end{cases} \quad \Longrightarrow \quad m = q = 0 \quad \Longrightarrow \quad y(x) = 0 \; \forall x$$

Quindi il minimo di $J(y)$ si ha in $J(0) = 2$.

Sembra sensato questo risultato? Bé, in realtà se ampliassimo leggermente l'insieme delle funzioni $y(x)$ prese in considerazione, aggiungendo alle funzioni $\mathcal{C}^1$ anche le funzioni $\mathcal{C}^1$ a tratti, troveremmo sicuramente valori più piccoli.

Ad esempio per una funzione di questo tipo:



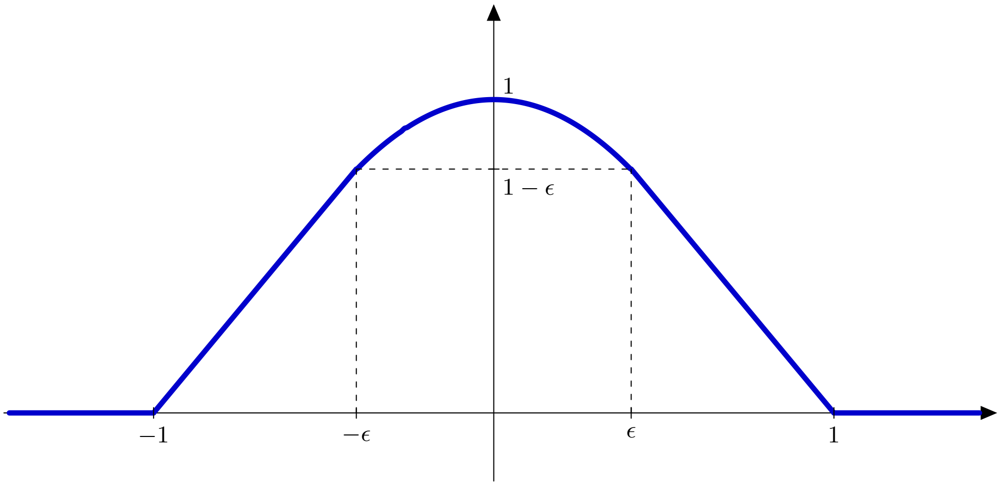
La funzione $J(y)$ assume il valore nullo: $J(y) = 0$, e questo non è sicuramente l'unico esempio.

Se ci pensiamo questo non è affatto strano: se ampliamo l'insieme in cui cerchiamo le funzioni, troveremo sicuramente un minimo che è minore o uguale a quello trovato in precedenza.

Il problema è che possiamo trovare funzioni $\mathcal{C}^1$ che risultano molto molto vicine a quella disegnata sopra, per cui quindi la funzione $J(y)$, che è continua, dovrebbe assumere valori vicini allo zero!

Vediamolo meglio:

Prendiamo una funzione di questo tipo:



Cerchiamo di scrivere la sua espressione:

La funzione sarà di tipo parabolico, ovvero del tipo $y(x) = ax^2 + c$, in cui $y(-\epsilon) = y(\epsilon) = 1 - \epsilon$, $y'(\pm\epsilon) = \mp 1$.

Calcoliamo: $y'(x) = 2ax \Longrightarrow y'(\epsilon) = 2a\epsilon$, quindi $y'(\epsilon) = -1 \Leftrightarrow 2a\epsilon = -1 \Leftrightarrow a = \frac{-1}{2\epsilon}$.

Sostituiamo quello che abbiamo trovato: $y(x) = \frac{-1}{2\epsilon}x^2 + c$.

Calcoliamo: $y(x) = -\frac{x^2}{2\epsilon} + c$, quindi $y(\epsilon) = 1 - \epsilon \Leftrightarrow c = 1 - \epsilon + \frac{\epsilon}{2} = 1 - \frac{\epsilon}{2}$.

Quindi la funzione è

$$y(x) = y(x) = -\frac{x^2}{2\epsilon} + 1 - \frac{\epsilon}{2}$$

$$y'(x) = \begin{cases} 1 & \text{se } x \in [-1, -\epsilon] \\ -1 & \text{se } x \in [\epsilon, 1] \\ -\frac{1}{\epsilon}x & \text{se } x \in]-\epsilon, \epsilon[\end{cases}$$

Possiamo calcolare il valore $J(y)$ in questa funzione:

$$J(y) = \int_{-1}^{1} \left(1 - y'(x)^2\right)^2 dx = \int_{-\epsilon}^{\epsilon} \left(1 - \frac{1}{\epsilon^2}x^2\right)^2 dx < \int_{-\epsilon}^{\epsilon} 1\, dx = 2\epsilon$$

Quindi abbiamo trovato una funzione $y(x) \in \mathcal{C}^1$ tale per cui $J(y)$ assume il valore 2ϵ “piccolo a piacere”, quindi evidentemente non è vero che 2 è valore di minimo per $J(y)$.

Capitolo 6

Ricerca Operativa - Parte 2

6.1 Decisioni in condizioni di incertezza

04/04/2014

In questo capitolo parleremo sempre di casi in cui le conseguenze sono di tipo monetario, in cui quindi la decisioni migliori saranno sempre quelle che ci fanno guadagnare di più. Questo per evitare di avere casi in cui la bontà delle decisioni dipende da un nostro giudizio personale che può essere diverso da individuo a individuo.

Speranza Matematica di X:

$$\mathbb{E}(X) = \sum x_k p_k \qquad p_k = \mathbb{P}(X = k)$$

$$\mathbb{E}(X) = \int_{\mathbb{R}} x f(x)\, dx \qquad f(x) = \text{ densità}$$

Quindi tra due guadagni aleatori sceglieremo sempre quello che ha speranza maggiore.

Esempio 6.1.1 Ci si propone di giocare dei soldi in un lancio di dadi, in cui puntando 1 euro, se esce testa si vincono 3 euro, se esce croce se ne vincono 0.

Il guadagno netto è: se esce testa 2 euro, se esce croce -1.

La speranza in questo caso è calcolabile: $\mathbb{E}(X) = 2\frac{1}{2} - 1\frac{1}{2} = \frac{1}{2} > 0$.

Quindi ci "conviene" giocare.

Nota: è chiaro che la conclusione di "convenienza" è valida fintanto che le somme di denaro restano basse rispetto ai propri fondi. Ad esempio, chiedere di giocare 1000 euro invece di 1 (e vincerne 3000 invece di 3) è esattamente lo stesso dal punto di vista della speranza matematica, ma è evidente che avremmo delle remore in più a giocarci uno stipendio piuttosto che un caffè.

Esempio 6.1.2 Un giornalaio deve acquistare n copie al costo unitario c che rivenderà al prezzo v.

In questo caso la variabile aleatoria è rappresentata dal numero di clienti che chiederanno questa specifica rivista: $X =$ numero di clienti.

Chiamiamo $p_h = \mathbb{P}(X = h)$ la probabilità che i clienti siano esattamente h, e calcoliamo il guadagno aleatorio associato alla decisione "n" (ovvero alla decisione di acquistare n copie):

$$G_n = \begin{cases} Xv - nc & \text{se } X \leq n \\ nv - nc & \text{se } X > n \end{cases}$$

Dobbiamo calcolare il valore atteso di G_n, così da poter capire quanto ci conviene:

$$\mathbb{E}(G_n) = v \left(\sum_{h=0}^{n} hp_h + n \sum_{h=n+1}^{\infty} p_h \right) - nc$$

Per scoprire quando la speranza è massima calcoliamo il valore successivo:

$$\mathbb{E}(G_{n+1}) = v \left(\sum_{h=0}^{n+1} hp_h + (n+1) \sum_{h=n+2}^{\infty} p_h \right) - (n+1)c$$

e facciamone la differenza

$$\mathbb{E}(G_{n+1}) - \mathbb{E}(G_n) = v \left((n+1)p_{n+1} - np_{n+1} + \sum_{h=n+2}^{\infty} p_h \right) - c =$$

$$= v \sum_{h=n+1}^{\infty} p_h - c = v \cdot \mathbb{P}(X > n) - c$$

Da cui ricaviamo la condizione

$$\mathbb{E}(G_{n+1}) \geq \mathbb{E}(G_n) \quad \Longleftrightarrow \quad \mathbb{P}(X > n) \geq \frac{c}{v}$$

Visto che per $n \longrightarrow +\infty$ la probabilità $\mathbb{P}(X > n) \longrightarrow 0$, ci sarà un determinato $\widehat{n}$ per cui $\mathbb{P}(X > \widehat{n}) \geq \frac{c}{v}$ ma $\mathbb{P}(X > (\widehat{n})) < \frac{c}{v}$. Questo sarà il numero di copie che ci conviene comprare.

Supponendo che la X sia una variabile di **Poisson**, le probabilità sono del tipo

$$\mathbb{P}(X = h) = \frac{\mu^h}{h!} e^{-\mu}$$

$$\mathbb{P}(X > n) = 1 - \sum_{h=0}^{n} \frac{\mu^h}{h!} e^{-\mu}$$

per cui, ad esempio, con $\mu = 10$ una tabella approssimativa delle probabilità risulta essere

n	7	8	9	10	11	12
$\mathbb{P}(X > n)$	0.78	0.67	0.54	0.42	0.30	0.21

Visto che la distribuzione di Poisson è piuttosto realistica per studi di questo tipo, potremmo basarci proprio su queste stime per fare gli acquisti dei giornali.

Il ragionamento fatto non è l'unico possibile, e forse nemmeno il migliore. Potremmo fare un conto di diverso tipo:

Il guadagno atteso dall'acquisto della prima copia di una rivista è $g_1 = -c+v\cdot\mathbb{P}(X \geq 1)$.

Il guadagno atteso dall'acquisto della seconda copia è $g_2 = -c+v\cdot\mathbb{P}(X \geq 2)$.

In generale il guadagno atteso dall'acquisto della n-esima copia è $g_n = -c+v\cdot\mathbb{P}(X \geq n)$.

Allora la decisione di quante copie comprare può essere basta su un ragionamento del tipo "ne acquisto n in modo che $g_n > 0$ e $g_{n+1} \leq 0$", così da massimizzare il guadagno atteso totale.

Esempio 6.1.3 Supponiamo di essere un orologiaio che acquista due tipi di orologi (A e B), che poi rivenderà a due prezzi diversi. Se non dovessimo vendere i pezzi, possiamo riportarli al fornitore, che non ci darà tutto il prezzo di partenza. La situazione può essere schematizzata:

	Vendita		Acquisto		Realizzo
A	v_A	$>$	c_A	$>$	r_A
B	v_B	$>$	c_B	$>$	r_B

Le probabilità di vendita per due clienti sono le seguenti:

$$\mathbb{P}(AA) = \frac{1}{4} \qquad \mathbb{P}(AB) = \frac{1}{2} \qquad \mathbb{P}(BB) = \frac{1}{4}$$

Poniamoci nel caso in cui i valori numerici sono i seguenti:

$$v_A = 30 \qquad c_A = 20 \qquad r_A = 8$$

$$v_B = 25 \qquad c_B = 17 \qquad r_B = 6$$

In questo caso possiamo schematizzare tutto con una tabella in cui le colonne rappresentano i guadagni rispetto ad una nostra decisione di quanti orologi comprare dal fornitore di quale tipo (ad esempio "$(1,2)$" rappresenta la decisione di prendere 1 orologio di tipo A e 2 di tipo B), e rispetto alle richieste dei nostri clienti (ad esempio "$0,2$" rappresenta l'evento per cui i due clienti vogliono entrambi l'orologio di tipo B):

	$(0,0)$	$(1,0)$	$(0,1)$	$(2,0)$	$(1,1)$	$(0,2)$	$(2,1)$	$(2,2)$	$(1,2)$	$\mathbb{P}$
$0,2$	0	-12	8	-24	-4	16	-16	-8	4	$\frac{1}{4}$
$1,1$	0	10	8	-2	18	-3	6	-5	7	$\frac{1}{2}$
$2,0$	0	10	-11	20	-1	-22	9	-2	-12	$\frac{1}{4}$
$\mathbb{E}$	0	4.5	3.25	-2	7.75	-3	$*$	$*$	$*$	

Per capire come si è riempita la tabella facciamo un esempio: se acquistiamo dal fornitore 1 orologio di tipo A e uno di tipo B (ovvero decisione "$(1, 1)$"), i guadagni sono i seguenti:

- Se i clienti vogliono entrambi un orologio di tipo B (situazione "$0, 2$") ne accontentiamo solo uno, quindi abbiamo pagato 20+17=37 al fornitore e ci riguadagniamo 25 vendendo l'orologio ad uno dei due. Ridando l'orologio invenduto al fornitore recuperiamo 8, quindi il guadagno totale è -37+25+8=-4, ovvero siamo in perdita.
- Se i clienti vogliono un orologio di un tipo A e uno di tipo B (situazione "$1, 1$"), abbiamo speso dal fornitore sempre lo stesso, 20+17=37, ma stavolta li abbiamo venduti entrambi, quindi il guadagno sarà -37+30+25=18.
- Se i clienti vogliono entrambi un orologio di tipo A (situazione "$2, 0$") ne accontento solo uno, quindi abbiamo pagato 20+17=37 al fornitore e ci riguadagniamo 30 vendendo l'orologio ad uno dei due. Ridando l'orologio invenduto al fornitore recuperiamo 6, quindi il guadagno totale è -37+30+6=-1, ovvero anche stavolta siamo in perdita.

Il valore atteso è poi calcolato ponendo i guadagni risultanti nelle rispettive probabilità.

Gli ultimi tre valori attesi non sono calcolati perché queste tre scelte sono "dominate" da altre, ovvero non ci conviene mai fare una di queste scelte poiché avremo sempre delle perdite che possiamo evitare prendendo una delle altre decisioni.

Questo perché i pezzi che rimangono invenduti in questo caso implicano perdite molto alte, quindi possiamo verificare nella tabella che in effetti per ognuna di queste tre colonne possiamo trovare una delle altre colonne che ha tutti e tre i valori più alti.

Quindi non ci conviene <u>mai</u> fare una di queste tre scelte!

La tabella che abbiamo scritto di chiama **schema in forma normale**, ovvero una tabella in cui sono contenute le decisioni da prendere (d_i), gli eventi possibili (e_i), le probabilità degli eventi (p_i), e i guadagni medi per ogni decisione (g_i):

	d_1	d_2	$\dots$	d_n	$\mathbb{P}$
e_1	x_{11}	x_{12}	$\cdots$	x_{1n}	p_1
e_2	x_{21}	x_{22}	$\cdots$	x_{2n}	p_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
e_m	x_{m1}	x_{m2}	$\cdots$	x_{mn}	p_m
$\mathbb{E}$	g_1	g_2		g_2	

07/04/2014

Esempio 6.1.4 Vediamo un esempio "medico". Si presenta una malattia, di cui abbiamo diversi dati:

Prevalenza della malattia: p (diffusione).

Test diagnostico: T =positivo (individuo malato), $\overline{T}$ =negativo (individuo sano/non malato).

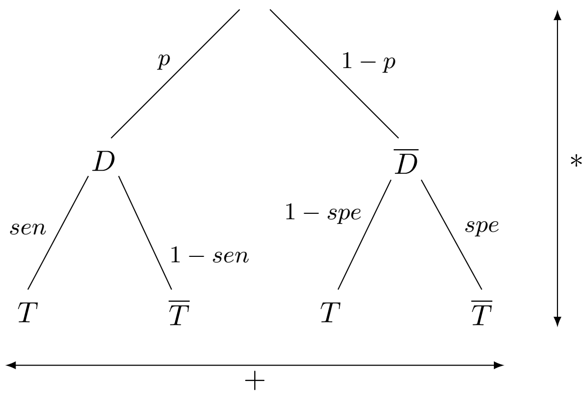
Sensibilità del test: $sen = \mathbb{P}(T|D)$ è la probabilità che il test dia esito positivo nel caso in cui il paziente sia malato.

Specificità del test: $spe = \mathbb{P}(\overline{T}, \overline{D})$ è la probabilità che il test dia esito negativo nel caso in cui il paziente sia sano.

Valore predittivo positivo: $V_{pp} = \mathbb{P}(D|T)$ è la probabilità che il il paziente sia malato a fronte di un test con esito positivo.

Valore predittivo negativo: $V_{pn} = \mathbb{P}(\overline{D}|\overline{T})$ è la probabilità che il il paziente sia sano a fronte di un test con esito negativo.

Stato di natura[1]:



Per capire come vengono calcolati alcuni dei valori facciamo un esempio delle quantità scritte sopra:

$$V_{pp} = \mathbb{P}(D|T) = \frac{\mathbb{P}(D|T)}{\mathbb{P}(T)} = \frac{\mathbb{P}(D)\mathbb{P}(T|D)}{\mathbb{P}(T)} = \frac{p \cdot sen}{p \cdot sen + (1-p)(1-spe)}$$

e così le altre.

Si presume che per casi derivanti da situazioni reali si abbia $p \simeq 0$, $sen \simeq 1$, $(1-spe) \simeq 0$.

Un test che risulta negativo non necessita di esser ripetuto perché il V_{pn} ha una buona probabilità di veridicità, infatti

$$V_{pn} = \mathbb{P}(\overline{D}|\overline{T}) = \frac{(1-p)spe}{\underbrace{p(1-sen)}_{\simeq 0} + (1-p)spe} \simeq 1$$

[1]In casi di questo tipo gli alberi grafici funzionano molto bene poiché le probabilità si sommano in orizzontale e si moltiplicano in verticale.

Capitolo 7

Distribuzioni in $\mathbb{R}$

7.1 Funzioni Test

08/04/2014

Come è ragionevole assegnare una probabilità? È una domanda che si pongono spesso i meteorologi, ma non solo loro.

In questo capitolo parleremo di questo argomento, ed in particolare delle *distribuzioni*, che si può trovare analizzato in maniera più erauriente nel volume *Theory of Distribution, a non-technical initroduction* di J. Ian Richards & H.K. Youn, Cambridge University

Se su un punto agisce una forza $F(t)$, nel tratto di tempo $[t_1, t_2]$ l'impulso è calcolabile come

$$\int_{t_1}^{t_2} F(t)\, dt = m\big(v(t_2) - v(t_1)\big)$$

Ma la forza è nulla in ogni istante diverso da quello in cui la palla sbatte su una sponda, supponiamo che sia $t_0 \in [t_1, t_2]$.

Se supponiamo che non ci siano forze esterne agenti, la palla si muove di moto rettilineo uniforme, ovvero $|v(t_1)| = |v(t_2)|$, ma se in questo tratto di tempo la palla ha sbattuto su una sponda la velocità dei due istanti sarà uguale in modulo, ma opposta in segno, ovvero

$$\int_{t_1}^{t_2} F(t)\, dt = m\big(v(t_2) - v(t_1)\big) = 2mv(t_2) \neq 0$$

seppur la funzione $F(t)$ sia nulla $\forall t \neq t_0$!

La media di una funzione f in un intervallo $[x_0 - \delta, x_0 + \delta] = I$ con una funzione test φ nulla al di fuori di I vale

$$\int_I f(x)\varphi(x)\, dx$$

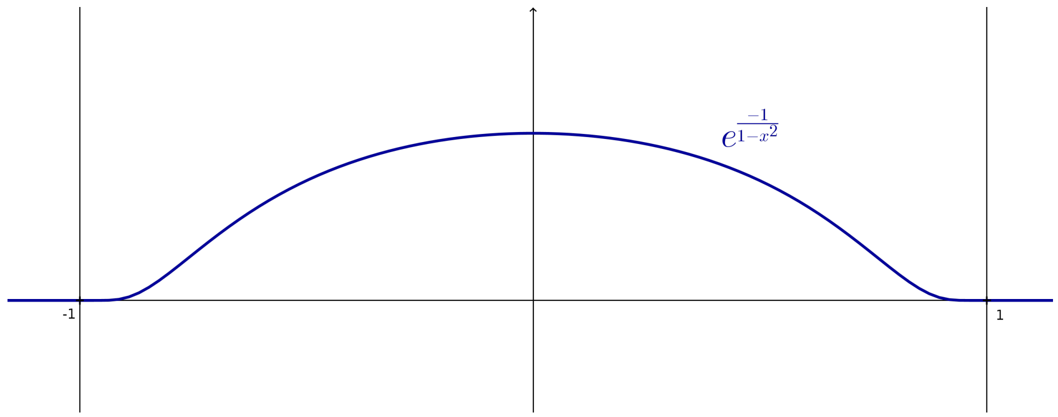
Abbiamo nominato poco fa le *funzioni test*, ovvero funzioni $\mathcal{C}^\infty$ a supporto compatto. Ma esistono veramente? Si, costruiamone una come esempio:

Prendiamo innanzitutto la funzione $\mathcal{C}^\infty$

$$g(t) = \begin{cases} e^{-\frac{1}{t}} & \text{se } t > 0 \\ 0 & \text{se } t \leq 0 \end{cases}$$

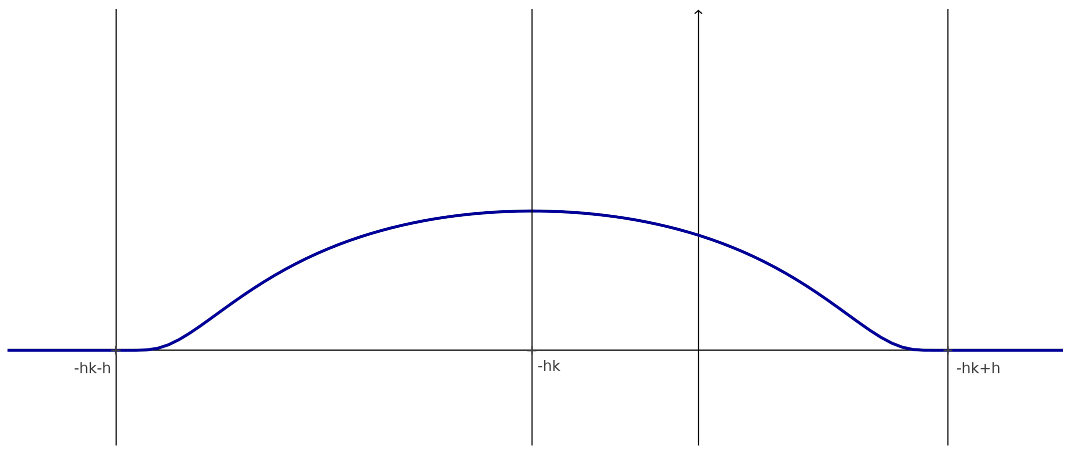
Allora $f(x) := g(1-x^2) \in \mathcal{C}^\infty$ è tale che $f(x) = 0 \;\; \forall |x| > 1$.

Questa funzione è proprio ciò che cercavamo, essendo $\mathcal{C}^\infty$ e con supporto $[-1, 1]$:



E se volessimo una funzione con supporto a nostra scelta, diciamo $[a, b]$?

Presi $h, k \in \mathbb{R}$, la funzione $F(x) := f\left(\frac{x}{h} + k\right)$ è una funzione "a campana" come la precedente (con stessa altezza), ma con supporto centrato nel punto di ascissa $-hk$ e di ampiezza $2|h|$, quindi la funzione è non nulla nell'insieme $[-hk-h, -hk+h]$.



Al variare di h e k possiamo quindi decidere a piacere il supporto di $F(x)$.

7.2 Definizioni di base

Ora che abbiamo chiarito l'esistenza e la definizione delle funzioni test iniziamo a parlare dell'argomento centrale del capitolo. Innanzitutto introduciamo la notazione $\mathcal{D}(\mathbb{R}) := \mathcal{C}_0^\infty(\mathbb{R}) = \{$ funzioni $\mathcal{C}^\infty$ a supporto compatto, definite su $\mathbb{R}$ $\}$, che sarà lo spazio (vettoriale) su cui costruiremo tutta la teoria.

Definizione 7.2.1 Una *distribuzione* è un funzionale lineare continuo da $\mathcal{D}(\mathbb{R})$ a $\mathbb{R}$. Lo spazio delle distribuzioni $\{T : \mathcal{D}(\mathbb{R}) \longrightarrow \mathbb{R} \mid T$ lineare e continua $\}$ è indicato con il simbolo $\mathcal{D}'(\mathbb{R})$.

Definizione 7.2.2 Ricordiamo che un funzionale lineare T si dice *continuo* se, presa una successione $\{\varphi_n\}$ in $\mathcal{D}(\mathbb{R})$ tale che $\varphi_n \longrightarrow 0$, vale $T(\varphi_n) \longrightarrow 0$.

11/04/2014

Osservazione Ricordiamo che presi $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ spazi normati, $f : X \longrightarrow Y$ è lineare e continua $\Longleftrightarrow \exists\, L > 0$ tale che $\forall x \in X$ vale $\| f(x) \|_Y \leq L \| x \|_Y$.

Teorema 7.2.1 *La funzione $T : \mathcal{D}(\mathbb{R}) \longrightarrow \mathbb{R}$ è lineare e continua $\Longleftrightarrow \forall K \subseteq \mathbb{R}$ compatto $\exists\, M > 0$ e $\exists\, m \in \mathbb{N}$ tali che $\forall \varphi \in \mathcal{D}(\mathbb{R})$ con supp$(\varphi) \subseteq K$ vale*

$$|T(\varphi)| \leq M \max_{\alpha \leq m} \max_{x \in K} \left|\varphi^{(\alpha)}(x)\right|$$

Dimostrazione. $\boxed{\Longleftarrow}$ Presa una successione (φ_n) di funzioni in $\mathcal{D}(\mathbb{R})$ convergente a 0, vorremmo dimostrare che anche $T(\varphi_n)$ converge a 0.

Grazie all'ipotesi di partenza questo si dimostra molto semplicemente:

$$|T(\varphi_n)| \leq \underbrace{M \max_{\alpha \leq m} \underbrace{\max_{x \in K} \left|\varphi_n^{(\alpha)}(x)\right|}_{\longrightarrow 0}}_{\longrightarrow 0} \longrightarrow 0 \quad \Longrightarrow \quad T \text{ è continua}$$

$\boxed{\Longrightarrow}$ Supponiamo che non valga la proprietà che vogliamo dimostrare, ovvero supponiamo $\exists\, K$ compatto per cui $\forall M > 0$, e $\forall m \in \mathbb{N}$ è sempre possibile trovare una $\varphi \in \mathcal{D}(\mathbb{R})$ con *supp*$(\varphi) \subseteq K$ tale che valga $|T(\varphi)| > M \max_{\alpha \leq m} \max_{x \in K} \left|\varphi^{(\alpha)}(x)\right|$.

Se $s \in \mathbb{N}$, prendiamo $M = m = s$. Per ciò che abbiamo supposto sappiamo che $\exists\, \varphi_s \in \mathcal{D}(\mathbb{R})$ con *supp*$(\varphi_s) \subseteq K$ per cui vale $|T(\varphi_s)| > s \max_{\alpha \leq m} \max_{x \in K} \left|\varphi_s^{(\alpha)}(x)\right|$.

Se definiamo

$$\psi_s := \frac{\varphi_s}{|T(\varphi_s)|}$$

risulta $|T(\psi_s)| = 1$.

Dunque se dimostriamo che la successione (ψ_s) converge a 0 in $\mathcal{D}(\mathbb{R})$ avremo trovato il nostro assurdo poiché risulterebbe che $T(\psi_s)$ non converge a 0, ovvero che T non è continua, che è in contrasto con l'ipotesi di partenza.

Le derivate delle funzioni ψ_s si calcolano facilmente

$$\psi_s^{(\alpha)} = \frac{\varphi_s^{(\alpha)}}{|T(\varphi_s)|}$$

e da queste ricaviamo che

$$1 = |T(\psi_s)| = \frac{1}{|T(\varphi_s)|}\,|T(\varphi_s)| > \frac{1}{|T(\varphi_s)|}\,s \max_{\alpha\leq m}\max_{x\in K}\left|\varphi_s^{(\alpha)}(x)\right| = s\max_{\alpha\leq m}\max_{x\in K}\left|\psi_s^{(\alpha)}(x)\right|$$

$$\Longrightarrow \qquad \max_{\alpha\leq m}\max_{x\in K}\left|\psi_s^{(\alpha)}(x)\right| < \frac{1}{s} \longrightarrow 0$$

□

Facciamo alcuni esempi di distribuzioni, per capire meglio di cosa si tratta.

Esempio 7.2.1 Presa $f : \mathbb{R} \longrightarrow \mathbb{R}$ continua, definiamo la seguente distribuzione:

$$\begin{array}{rcl} T_f : \mathcal{D}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & \int_{\mathbb{R}} f(x)\varphi(x) \end{array}$$

questa è una buona definizione, possiamo infatti provare che vale la proprietà scritta nel teorema precedente: fissato un compatto $K \subseteq \mathbb{R}$, se $supp(\varphi) \subseteq K$ abbiamo che

$$|T_f(\varphi)| \leq \int_K f(x)\varphi(x) \leq \overbrace{\int_K |f(x)|}^{M}\ \max_{x\in K}|\varphi(x)|$$

In realtà anziché T sia continua è sufficiente che f sia sommabile su ogni compatto di $\mathbb{R}$, ovvero $f \in L^1_{loc}(\mathbb{R}) = \{g : \mathbb{R} \to \mathbb{R} \mid g \text{ misurabile}, \forall K \text{ compatto vale } \int_K |g| \in \mathbb{R}\}$.

Notazione: d'ora in avanti invece di scrivere $T(\varphi)$ scriveremo $< T, \varphi >$.

Sorge una domanda: le distribuzioni saranno tutte di "tipo funzione", ovvero derivanti da una funzione f come nell'esempio precedente? In realtà <u>no</u>. La distribuzione definita come

$$< \delta, \varphi > := \varphi(0)$$

è lineare e per essa vale la proprietà del teorema, infatti

$$| < \delta, \varphi > | = |\varphi(0)| \leq \max_{x\in K}|\varphi(x)|$$

ma non può essere di tipo funzione! Dimostriamolo:

Se per assurdo esistesse $f \in L^1_{loc}(\mathbb{R})$ tale che $\forall \varphi \in \mathcal{D}(\mathbb{R})$ vale $\int_{\mathbb{R}} f\varphi = \varphi(0)$, allora potremmo prendere una φ con $supp(\varphi) \subseteq \mathbb{R}\backslash\{0\}$, e risulterebbe $\int_{\mathbb{R}} f\varphi = 0$.

Allora f sarebbe nulla quasi ovunque in $\mathbb{R}\backslash\{0\}$, quindi nulla quasi ovunque in $\mathbb{R}$.

Questo implica che $\varphi(0) = \int_{\mathbb{R}} f\varphi = 0 \quad \forall\varphi \in \mathcal{D}(\mathbb{R})$, il che è assurdo.

Quindi δ non è una distribuzione di tipo funzione.

Proprietà:

- Prese $S, T \in \mathcal{D}'(\mathbb{R})$ e $\lambda \in \mathbb{R}$, allora $S + T \in \mathcal{D}'(\mathbb{R})$ e $\lambda T \in \mathcal{D}'(\mathbb{R})$.

- Se $T \in \mathcal{D}'(\mathbb{R})$ e $\alpha \in \mathbb{R}$, allora $T(x-\alpha)$ è una distribuzione che verifica

$$< T(x-\alpha), \varphi(x) > = < T(x), \varphi(x+\alpha) >$$

Oppure possiamo scrivere

$$< T(x-\alpha), \varphi(x) > = \int f(x-\alpha)\varphi(x)\, dx \doteq \int f(y)\varphi(y-a)\, dy = < T(y), \varphi(y-a) >$$

Osservazione Grazie a queste proprietà possiamo dimostrare altre utili uguaglianze:

$$< \delta(x), \varphi(x+a) > = < \delta(x-a), \varphi(x) > = \varphi(a)$$

$$< T(ax), \varphi(x) > = \frac{1}{|a|} < T(x), \varphi\left(\frac{x}{a}\right) >$$

7.3 Derivata di una distribuzione

14/04/2014

Se $T \in \mathcal{D}'(\mathbb{R})$, T' è la distribuzione definita per $\varphi \in \mathcal{D}(\mathbb{R})$ da

$$< T', \varphi >= - < T, \varphi' >$$

Che è una buona definizione perché è rispettata la proprietà $| < T', \varphi > | = | < T, \varphi' > |$, e se supponiamo $\varphi \subseteq K$, allora vale

$$| < T, \varphi' > | \leq M \max_{\alpha \leq m+1} \max_{x \in K} \left|\varphi^{(\alpha)}(x)\right|$$

Invece se $T = T_f$, con $f \in \mathcal{C}^1(\mathbb{R})$, chi è T'_f? Vorremmo che fosse $T_{f'}$, e in effetti è così perché

$$< T', \varphi >= - < T, \varphi' >= - \int_{-\infty}^{+\infty} f\varphi' =$$

$$= -\cancel{\Big[f\varphi\Big]_{-\infty}^{+\infty}} + \int_{-\infty}^{+\infty} f'\varphi = \int_{-\infty}^{+\infty} f'\varphi$$

Da notare che in realtà la funzione φ ha un supporto compatto, quindi tutti gli integrali non saranno su tutto $\mathbb{R}$ ma limitati in un intervallo compatto $[a, b]$.

Cosa succede se $T = T_f$ con f non derivabile?
Supponiamo f non derivabile ma continua in x_0.

$$< T', \varphi >= - < T, \varphi' >= - \int_{-\infty}^{+\infty} f\varphi' = - \int_{-\infty}^{x_0} f\varphi' - \int_{x_0}^{+\infty} f\varphi' =$$

$$= -\Big[f\varphi\Big]_{-\infty}^{x_0} + \int_{-\infty}^{x_0} f'\varphi - \Big[f\varphi\Big]_{x_0}^{+\infty} + \int_{x_0}^{+\infty} f'\varphi =$$

$$= -f(x_0)\varphi(x_0) + \int_{-\infty}^{+\infty} f'\varphi + f(x_0)\varphi(x_0) = \int_{-\infty}^{+\infty} f'\varphi$$

Quindi non cambia niente.

Se invece f ha proprio un salto di discontinuità in x_0 ma è continua e derivabile in tutti gli altri punti avremo

$$< T', \varphi >= - < T, \varphi' >= - \int_{-\infty}^{+\infty} f\varphi' = - \int_{-\infty}^{x_0} f\varphi' - \int_{x_0}^{+\infty} f\varphi' =$$

$$= \lim_{x \longrightarrow x_0^-} \left(-\Big[f\varphi\Big]_{-\infty}^{x} \right) + \int_{-\infty}^{x_0} f'\varphi + \lim_{x \longrightarrow x_0^+} \left(-\Big[f\varphi\Big]_{x}^{+\infty} \right) + \int_{x_0}^{+\infty} f'\varphi =$$

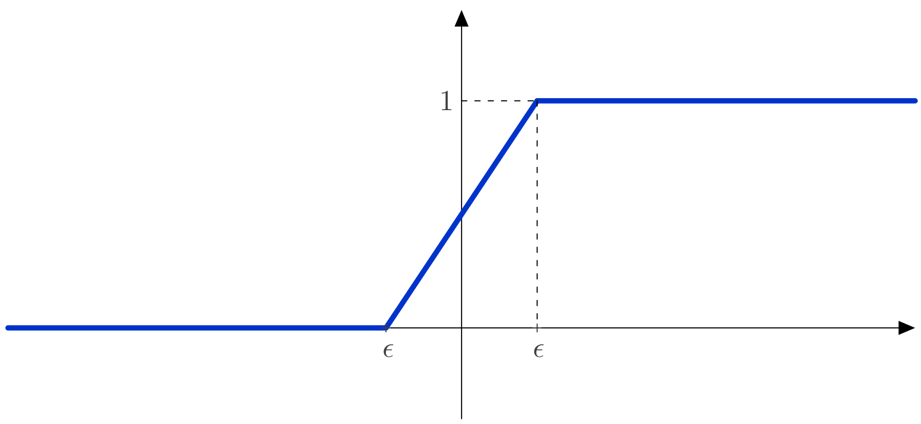
$$= -l_1\varphi(x_0) + \int_{-\infty}^{+\infty} f'\varphi + l_2\varphi(x_0) = \int_{-\infty}^{+\infty} f'\varphi =< T_{f'}, \varphi > + \big(l_2 - l_1\big)\varphi(x_0) =$$

$$= < T_{f'} + \big(l_2 - l_1\big)\delta(x - x_0), \varphi >$$

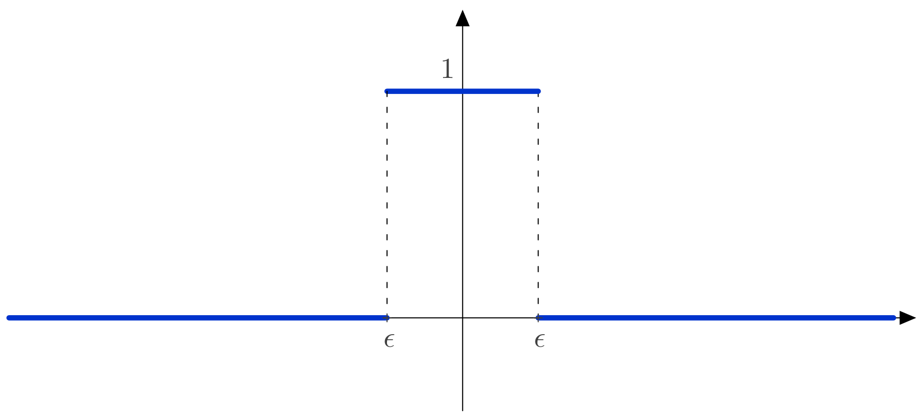
Presa la funzione $H(x) = \chi_{[0,+\infty[}$ avremo $(T_H)' = \delta$, che ci descrive meglio il comportamento della funzione rispetto ad una derivata scritta come

$$H'(x) = \begin{cases} 0 & \text{se } x < 0 \\ 0 & \text{se } x > 0 \\ \nexists & \text{se } x = 0 \end{cases}$$

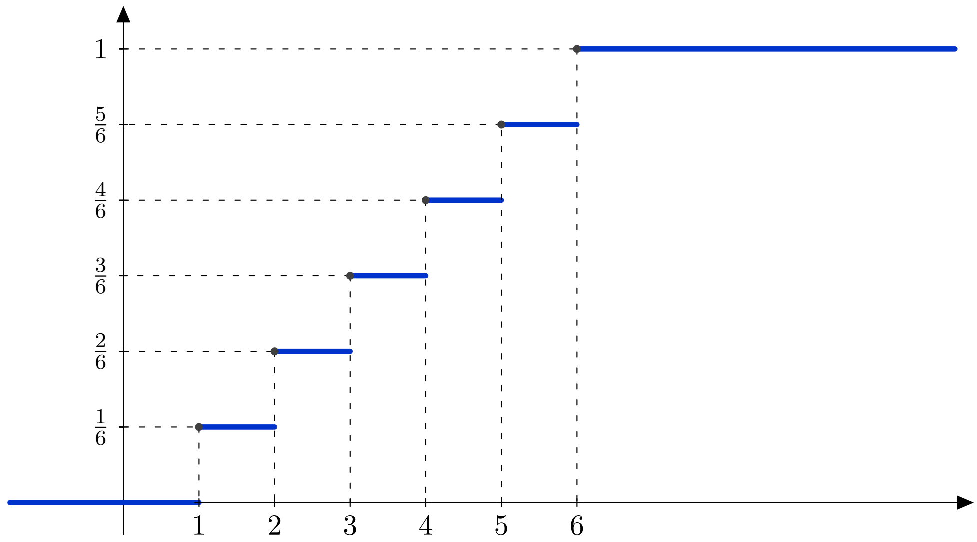
Possiamo infatti definire una funzione $H_\epsilon(x)$ con un comportamento di questo tipo:



con derivata che risulta



Con una funzione più complessa del tipo $F_X(x) = \mathbb{P}(X \leq x)$ risulta invece



Con derivata (che non disegniamo) di questa forma:

$$F'_X = \sum_{k=1}^{6} \frac{1}{6}\delta(x-k)$$

Descriviamo quindi le **Regole di derivazione** per le distribuzioni:

- $(S+T)' = S' + T'$

  Dimostrazione.

  $$< (S+T)', \varphi >= - < S+T, \varphi' >= - < S, \varphi' > - < T, \varphi' >=$$
  $$=< S'\varphi > + < T', \varphi >=< S' + T', \varphi >$$

  □

- $(aT)' = aT'$

  Dimostrazione.

  $$< (aT)', \varphi >= - < aT, \varphi' >= - < T, a\varphi' >= - < T, (a\varphi)' >=$$
  $$=< T', a\varphi >=< aT', \varphi >$$

  □

- $(T(ax))' = aT'(ax)$

  Dimostrazione.

$$< (Tax))', \varphi >= - < T(ax), \varphi'(x) >= - < T(x), \varphi'\left(\frac{x}{a}\right) > \cdot \frac{1}{a} =$$

$$= - < T(x), \frac{1}{a}\varphi'\left(\frac{x}{a}\right) >= - < T(x), \left(\varphi\left(\frac{x}{a}\right)\right)' >=< T'(x), \varphi\left(\frac{x}{a}\right) > \cdot \frac{1}{a} \cdot a =$$
$$= a < T'(ax), \varphi >$$

$\square$

- $(T(x-a))' = T'(x-a)$

  Dimostrazione.

$$< T(x-a)', \varphi >= - < T(x-a), \varphi'(x) >= - < T(x), \varphi'(x+a) >=$$

$$= - < T(x), \big(\varphi(x+a)\big)' >=< T'(x), \varphi(x+a) >=< T'(x-a), \varphi >$$

$\square$

- Se $g \in \mathcal{C}^\infty(\mathbb{R})$ allora $(gT)' = g'T + gT'$

  Dimostrazione.

$$< gT', \varphi >=< T', g\varphi >= - < T, (g\varphi)' >= - < T, g'\varphi + g\varphi' >=$$

$$= - < T, g'\varphi > - < T, g\varphi' >= - < g'T, \varphi > - < gT, \varphi' >=$$
$$= - < g'T, \varphi > + < (gT)', \varphi >$$

$\square$

7.4 Convergenza in $\mathcal{D}'(\mathbb{R})$

Definizione 7.4.1 Una successione (T_k) in $\mathcal{D}'(\mathbb{R})$ converge a $T \in \mathcal{D}'(\mathbb{R})$ se per ogni $\varphi \in \mathcal{D}(\mathbb{R})$ vale $< T_k, \varphi > \longrightarrow < T, \varphi >$.

Proposizione 7.4.1 *Prese due successioni $(T_k), (S_k)$ convergenti rispettivamente a T, S e $a, x \in \mathbb{R}$, valgono le seguenti relazioni:*

1. $T_k + S_k \longrightarrow T + S$
2. $aT_k \longrightarrow aT$

3. $T_k(ax) \longrightarrow T(ax)$

4. $T_k(x-a) \longrightarrow T(x-a)$

5. $gT_k \longrightarrow gT$

6. $T'_k \longrightarrow T'$

Cosa accade se la successione è del tipo $T_k = T_{f_k}$, con $f_k \longrightarrow f$ in qualche senso[1]? In alcuni casi sarà vero che

$$< T_{f_k}, \varphi >= \int f_k\varphi \longrightarrow \int f\varphi$$

ma quando? In realtà basta poco: se $f_k \longrightarrow f$ puntualmente e $\forall D$ compatto vale che $\sup_{x\in D, k\in\mathbb{N}} |f_k(x)| \in \mathbb{R}$ e già abbiamo vinto.

7.5 Approssimazione della distribuzione δ

Teorema 7.5.1 *Sia f sommabile in $\mathbb{R}$, con $\int_{\mathbb{R}} f = 1$ e chiamiamo $\forall a > 0$ $f_a(x) = af(ax)$. Allora $\lim_{a\longrightarrow+\infty} f_a = \delta$ in $\mathcal{D}'(\mathbb{R})$.*

Dimostrazione. Vogliamo dimostrare che $\forall\varphi \in \mathcal{D}(\mathbb{R})$ vale

$$\lim_{a\longrightarrow+\infty} \int_{\mathbb{R}} f_a(x)\varphi(x)\, dx - \varphi(0) = 0$$

Innanzitutto manipoliamo l'espressione appena scritta:

$$\int_{\mathbb{R}} f_a(x)\varphi(x)\, dx - \varphi(0) \doteq \int_{\mathbb{R}} f(u) \left(\varphi\left(\frac{u}{a}\right) - \varphi(0)\right)\, du$$

operando il cambio di variabile $u = ax$.

Preso $\epsilon > 0$ sappiamo che $\exists\, M > 0$ tale che $\int_{|u|>M} |f| < \epsilon$, dunque spezziamo l'integrale trovato prima in due parti:

$$\underbrace{\int_{|u|\leq M} f(u) \left(\varphi\left(\frac{u}{a}\right) - \varphi(0)\right)\, du}_{A} + \underbrace{\int_{|u|>M} f(u) \left(\varphi\left(\frac{u}{a}\right) - \varphi(0)\right)\, du}_{B}$$

e ragioniamo sui due integrali trovati.

$$|A| \leq \int_{|u|\leq M} |f(u)| \left|\varphi\left(\frac{u}{a}\right) - \varphi(0)\right|\, du$$

[1] L'espressione "in qualche senso" è doverosa poiché la convergenza di funzioni può essere di molti tipi. Qui stiamo semplicemente richiedendo che si abbia una convergenza di qualunque tipo.

$\exists\, \gamma$ tale che $|x| < \gamma \Longrightarrow |\varphi(x) - \varphi(0)| < \epsilon$.
Prendiamo a in modo che $\frac{M}{a} < \gamma$, allora

$$|A| \leq \epsilon \int_{|u| \leq M} |f| \leq \int_{\mathbb{R}} |f| = \epsilon \cdot C$$

Adesso l'altro:

$$|B| \leq \int_{|u| > M} |f| \; 2 \max |\varphi| \; du < \epsilon \cdot 2 \max |\varphi|$$

Possiamo quindi concludere tornando all'integrale di partenza:

$$\left| \int_{\mathbb{R}} f_a(x) \varphi(x) \; dx - \varphi(0) \right| < \epsilon \left(C + 2 \max |\varphi| \right)$$

□

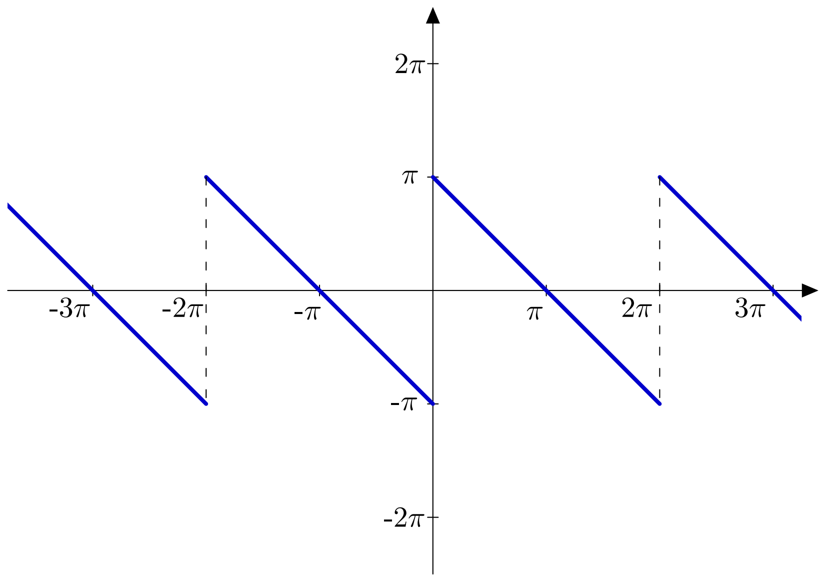
7.6 Serie di Fourier di distribuzioni

15/04/2013

Esattamente come si fa con le successioni numeriche, è possibile considerare serie di distribuzioni, definite come

$$\sum_{k=1}^{\infty} T_k := \lim_{n \longrightarrow +\infty} S_n \left(= \sum_{k=1}^{n} T_k \right) = S$$

Esempio 7.6.1 La funzione f descritta come nel grafico



è 2π-periodica, quindi possiamo scrivere la *serie di Fourier* della funzione f:

$$S(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \big(a_n \cos(nx) + b_n \sin(nx) \big)$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx \qquad\qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$$

Nel nostro caso, per come è fatta $f(x)$, tutti gli a_n risultano nulli, quindi dobbiamo solo analizzare i b_n:

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx = \frac{2}{\pi}\int_0^{\pi}(\pi - x)\sin(nx)\,dx =$$

$$= \frac{2}{\pi}\left(\left[-\frac{1}{n}\cos(nx)\cdot(\pi - x)\right]_{x=0}^{x=\pi} - \frac{1}{n}\int_0^{\pi}\cos(nx)\,dx\right) = \frac{2}{n}$$

Quindi la serie $S(x)$ risulta essere

$$S(x) = \sum_{n=1}^{\infty}\frac{2}{n}\sin(nx)$$

Lo spazio in cui si fanno solitamente tutte queste considerazioni è lo spazio di funzioni L^2, ma che legame c'è tra questo e il nostro spazio $\mathcal{D}'$?

Lemma *La convergenza L^2 implica la convergenza in $\mathcal{D}'$.*

Dimostrazione. Se $f_n \longrightarrow f$ in L^2 vale

$$\left|\int (f_n - f)\,\varphi\right| \leq \int |f_n - f|\cdot|\varphi| \leq \| f_n - f \|_2 \cdot \| \varphi \|_2$$

□

In $\mathcal{D}'(\mathbb{R})$ possiamo scrivere

$$T_f = f(x) = \sum_{n=1}^{\infty}\frac{2}{n}\sin(nx)$$

Ma abbiamo detto che $S_n \longrightarrow S \Longrightarrow S_n' \longrightarrow S'$, e questo ci permette di poter derivare ogni serie convergente termine a termine ed ottenere ancora una serie convergente, quindi

$$T_f' = \sum_{n=1}^{\infty} 2\cos(nx)$$

Ma l'espressione che sappiamo calcolare di T_f', visto che $f'(x) = -1$ "quasi ovunque", risulta

$$T_f' = -1 + 2\pi\sum_{k\in\mathbb{Z}}\delta(x - 2k\pi)$$

$$\Longrightarrow < T_f', \varphi >= \sum_{\substack{k\in\mathbb{Z}\\ 2k\pi\in Supp(\varphi)}} 2\pi\varphi(2k\pi) - \int_{\mathbb{R}}\varphi$$

Questa serie non converge in nessun senso "tradizionale", mentre pensata come distribuzione sappiamo che converge perché abbiamo un teorema che lo assicura!

Nessuno ci vieta di continuare a derivare, quindi in teoria potremmo continuare ancora per trovare

$$\sum_{n=1}^{\infty} -2n\sin(nx) = 2\pi \sum_{k\in\mathbb{Z}} \delta'(x-2k\pi)$$

e così via...

7.7 Distribuzioni particolari

Studiamo adesso le *pseudo-funzioni* $\frac{1}{x^n}$, con $n \in \mathbb{N}$, viste come distribuzioni.

$\boxed{n=1}$ Abbiamo $\frac{1}{x}$, e definiamo

$$< \frac{1}{x}, \varphi(x) >:= \lim_{\epsilon\to 0^+} \int_{|x|>\epsilon} \frac{1}{x}\varphi(x)\, dx = \circledast$$

perché questa definizione sia sensata dobbiamo dimostrare che il limite esiste e che quella scritta sia veramente una distribuzione. Vediamo il limite:

$$\circledast = \lim_{\epsilon\to 0^+} \left(\int_{-\infty}^{-\epsilon} \frac{1}{x}\varphi(x)\, dx + \int_{\epsilon}^{+\infty} \frac{1}{x}\varphi(x)\, dx \right) =$$

facciamo una sostituzione $y = -x$ per riscrivere meglio il primo integrale

$$= \lim_{\epsilon\to 0^+} \left(\int_{+\infty}^{\epsilon} \frac{1}{y}\varphi(-y)\, dy + \int_{\epsilon}^{+\infty} \frac{1}{x}\varphi(x)\, dx \right) =$$

$$= \lim_{\epsilon\to 0^+} \left(\int_{\epsilon}^{+\infty} -\frac{1}{y}\varphi(-y)\, dy + \int_{\epsilon}^{+\infty} \frac{1}{x}\varphi(x)\, dx \right) =$$

il fatto che all'interno dell'integrale ci sia la variabile y o x non cambia niente, quindi possiamo riaccorpare gli integrali e scriverli con la sola variabile x:

$$= \int_{\epsilon}^{+\infty} \frac{\varphi(x)-\varphi(-x)}{x}\, dx = \int_{0}^{+\infty} \frac{\varphi(x)-\varphi(-x)}{x}\, dx$$

Quindi il limite esiste.

Osservazione Se $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ e $\varphi(0) = 0$ allora $\frac{\varphi(x)}{x} \in \mathcal{C}^\infty(\mathbb{R})$.

$\boxed{n>1}$ Conosciamo le regole di derivazione. Allora da queste ricaviamo quello che ci serve:

$$\frac{d}{dx}x^{-n} = -nx^{-n-1} = \frac{-n}{x^{n+1}} \Longrightarrow \frac{1}{x^{n+1}} = -\frac{1}{n}\frac{d}{dx}\left(\frac{1}{x^n}\right)$$

Quindi abbiamo definito la distribuzione di x^{n+1} tramite un ragionamento induttivo dai casi precedenti, il che è possibile perché abbiamo definito esplicitamente il caso $n = 1$.

Teorema 7.7.1

$$x \cdot \frac{1}{x^{n+1}} = \frac{1}{x^n}$$

Dimostrazione. Per induzione:

$\boxed{n=1}$ Dobbiamo dimostrare che $x\frac{1}{x} = 1$:

$$< x\frac{1}{x}, \varphi(x) > = < \frac{1}{x}, x\varphi(x) > = \lim_{\epsilon \to 0^+} \int_{|x| \geq \epsilon} \frac{1}{x} x\varphi(x)\, dx = \int_{\mathbb{R}} \varphi(x)\, dx \ = < 1, \varphi(x) >$$

$\boxed{n+1}$ Di nuovo:

$$< x\frac{1}{x^{n+1}}, \varphi(x) > = < \frac{1}{x^{n+1}}, x\varphi(x) > = -\frac{1}{n} < \frac{d}{dx}\left(\frac{1}{x^n}\right), x\varphi(x) > =$$

$$= \frac{1}{n} < \frac{1}{x^n}, \frac{d}{dx}(x\varphi(x)) > = \frac{1}{n} < \frac{1}{x^n}, \varphi(x) + x\varphi'(x) > =$$

$$= \frac{1}{n} < \frac{1}{x^n}, \varphi(x) > + \frac{1}{n} < x\frac{1}{x^n}, \varphi'(x) > \doteq \frac{1}{n} < \frac{1}{x^n}, \varphi(x) > + \frac{1}{n} < \frac{1}{x^{n-1}}, \varphi'(x) > =$$

$$= \frac{1}{n} < \frac{1}{x^n}, \varphi(x) > - \frac{1}{n} < \frac{d}{dx}\frac{1}{x^{n-1}}, \varphi(x) > =$$

$$= \frac{1}{n} < \frac{1}{x^n}, \varphi(x) > + \frac{n-1}{n} < \frac{1}{x^n}, \varphi(x) > = < \frac{1}{x^n}, \varphi(x) >$$

□

Proposizione 7.7.1 *Prese T distribuzione, φ funzione test e $a \in \mathbb{R}$, valgono le seguenti uguaglianze:*

$$< T(ax), \varphi(x) > = < T(x), \frac{1}{|a|}\varphi\left(\frac{x}{a}\right) >$$

$$< T(-x), \varphi(x) > = < T(x), \varphi(-x) >$$

Esattamente come si fa per le funzioni reali, possiamo definire la parità di una distribuzione.

Definizione 7.7.1 T è *pari* se $T(-x) = T(x)$, mentre è *dispari* se $T(-x) = -T(x)$.

Proposizione 7.7.2 *Se T è pari, T' è dispari, e viceversa.*

Dimostrazione. Supponiamo T pari:

$$< T'(-x), \varphi(x) > = < T'(x), \varphi(-x) > = \ - < T(x), \varphi'(-x) > =$$

$$= < T(-x), \varphi'(-x) > = < T(x), \varphi'(x) > = < -T'(x), \varphi(x) >$$

$\Longrightarrow$ T' è dispari: $T'(-x) = -T'(x)$.

L'altra implicazione si dimostra allo stesso modo.

□

Proposizione 7.7.3 *$\frac{1}{x^n}$ è pari (dispari) se n è pari (dispari).*

Dimostrazione. $\frac{1}{x}$ è dispari:

$$< \frac{1}{-x}, \varphi(x) > = < \frac{1}{x}, \varphi(-x) > = \lim_{\epsilon \to 0^+} \int_{|x| \geq \epsilon} \frac{1}{x} \varphi(-x)\, dx =$$

applichiamo il cambiamento di variabile $y = -x$:

$$= \lim_{\epsilon \to 0^+} \int_{|y| \geq \epsilon} \frac{1}{-y} \varphi(y)\, dy = - \lim_{\epsilon \to 0^+} \int_{|y| \geq \epsilon} \frac{1}{y} \varphi(y)\, dy = - < \frac{1}{x}, \varphi(x) >$$

Dopodiché il resto si ha perché moltiplichiamo funzioni pari/dispari $\frac{1}{x^n}$ per la funzione dispari $\frac{1}{x}$, ottenendo così funzioni dispari/pari.

□

Definizione 7.7.2 T è *periodica* (di periodo a) se $T(x - a) = T(x)$

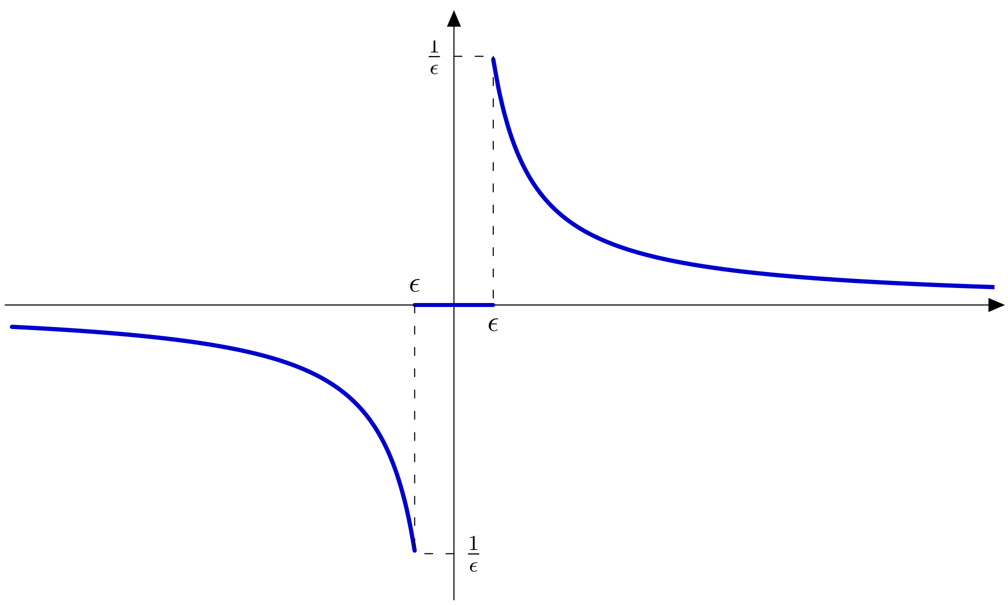
Definizione 7.7.3 T è *positiva* se $\forall \varphi \geq 0$ vale $< T, \varphi > \geq 0$.

Proposizione 7.7.4 *Prese T distribuzione, φ funzione test e $a \in \mathbb{R}$, vale*

$$< T(x - a), \varphi(x) > = < T(x), \varphi(x + a) >$$

Esempio 7.7.1 Definiamo la seguente funzione

$$\left(\frac{1}{x}\right)_\epsilon = \begin{cases} \frac{1}{x} & \text{se } |x| \geq \epsilon \\ 0 & \text{se } |x| < \epsilon \end{cases}$$



In $\mathcal{D}'(\mathbb{R})$ abbiamo che

$$\frac{1}{x} = \lim_{\epsilon\to 0^+}\left(\frac{1}{x}\right)_\epsilon$$

Allora la distribuzione $\frac{1}{x^2}$ possiamo scriverla come

$$\frac{1}{x^2} = -\frac{d}{dx}\frac{1}{x} = -\frac{d}{dx}\lim_{\epsilon\to 0^+}\left(\frac{1}{x}\right)_\epsilon = -\lim_{\epsilon\to 0^+}\frac{d}{dx}\left(\frac{1}{x}\right)_\epsilon =$$

$$= -\lim_{\epsilon\to 0^+}\left(\left(-\frac{1}{x^2}\right)_\epsilon + \frac{1}{\epsilon}\left(\delta_{-\epsilon}+\delta_\epsilon\right)\right)$$

Ma per i valori di δ vale

$$\lim_{\epsilon\to 0}\frac{\delta_\epsilon + \delta_{-\epsilon} - 2\delta_0}{\epsilon} = 0$$

questo perché

$$\lim_{\epsilon\to 0}\frac{\varphi(\epsilon)+\varphi(-\epsilon)-2\varphi(0)}{\epsilon} \overset{\text{H}}{=} \lim_{\epsilon\to 0}\frac{\varphi'(\epsilon)-\varphi'(-\epsilon)}{1} = 0$$

$$\implies \quad < \frac{\delta_\epsilon + \delta_{-\epsilon} - 2\delta_0}{\epsilon}, \varphi > = \frac{1}{\epsilon}\left(\varphi(\epsilon)+\varphi(-\epsilon)-2\varphi(0)\right)$$

Il che ci porta

$$\frac{1}{x^2} = \lim_{\epsilon\to 0}\left(\left(\frac{1}{x^2}\right)_\epsilon - \frac{2\delta_0}{\epsilon}\right)$$

Osservazione $\frac{1}{x^2}$ non è positiva.

Dimostrazione. Presa una funzione funzione test $\varphi \geq 0$ con $\max\varphi = \varphi(0)$ e $\varphi(x) = 0$ se $|x| \geq M$, vale

$$< \frac{1}{x^2}, \varphi(x) > = \lim_{\epsilon\to 0^+}\left(\int_\epsilon^{+\infty}\frac{\varphi(x)}{x^2}\,dx - \frac{2}{\epsilon}\varphi(0)\right) \leq \lim_{\epsilon\to 0^+}\left(2\varphi(0)\int_\epsilon^M \frac{1}{x^2}\,dx - \frac{2}{\epsilon}\varphi(0)\right) \leq$$

$$\leq \lim_{\epsilon\to 0^+}\left(2\varphi(0)\left(-\frac{1}{\epsilon} - \frac{1}{M} + \frac{1}{\epsilon}\right)\right) = \frac{2\varphi(0)}{M}$$

□

7.8 Trasformata di Fourier

28/04/2014

Definizione 7.8.1 Se $f \in L'(\mathbb{R})$, definiamo la *trasformata di Fourier di* f la seguente funzione

$$\widehat{f}(t) := \int_{\mathbb{R}} e^{-2\pi i t x} f(x)\,dx$$

e l'*antitrasformata di Fourier di* f la funzione

$$\overset{\vee}{f}(t) := \int_{\mathbb{R}} e^{2\pi i t x} f(x)\,dx$$

Osservazione Se $f \in L'(\mathbb{R})$, allora $\int_{\mathbb{R}} |f| \in \mathbb{R}$.
Inoltre, se $t \in \mathbb{R}$, allora $|f(x)| = |e^{-2\pi i t x} f(x)|$.

Teorema 7.8.1 (di Riemann-Lebesgue) *Se $f \in L'(\mathbb{R})$, allora $\widehat{f}$ è uniformemente continua e vale*

$$\lim_{|t| \to +\infty} \widehat{f}(t) = 0$$

Proprietà della trasformata di Fourier:

1. $(af + bg)^\wedge = a\widehat{f} + b\widehat{g}$ (*linearità*)
2. $\widehat{f'}(t) = 2\pi i t \cdot \widehat{f}(t)$ (f derivabile con $f' \in L'$)
3. Se $f \in L'$, allora $xf(x) \in L'$, e valgono
$$\big(xf(x)\big)^\vee(t) = \frac{1}{2\pi i}\left(\overset{\vee}{f}\right)'(t) \qquad\qquad -\big(xf(x)\big)^\wedge(t) = -\frac{1}{2\pi i}\left(\widehat{f}\right)'(t)$$
4. $\big(f(x+a)\big)^\wedge(t) = e^{2\pi i a t} \cdot \widehat{f}(t)$
5. $\big(e^{2\pi i a x} \cdot f(x)\big)^\vee(t) = \overset{\vee}{f}(t+a)$
$\big(e^{2\pi i a x} \cdot f(x)\big)^\wedge(t) = \widehat{f}(t-a)$
6. $\big(f(ax)\big)^\wedge(t) = \frac{1}{|a|}\widehat{f}\left(\frac{t}{a}\right)$
7. $(f \cdot g)^\wedge(t) = \widehat{f}(t) \cdot \widehat{g}(t)$, dove la moltiplicazione è definita come
$$(f \cdot g)(x) = \int_{\mathbb{R}} f(x-y)g(y)\, dy$$
8. $\left(\overline{f(-x)}\right)^\wedge(t) = \overline{\widehat{f}}(t)$
9. $\int \widehat{f} g = \int f \widehat{g}$

Definizione 7.8.2 Diciamo che f è *a decrescenza rapida* (funzione *DR*) se vale

$$\lim_{|x| \to \infty} x^n f(x) = 0 \quad \forall n \in \mathbb{N}$$

viceversa, diciamo che g è *a crescenza lenta* (funzione *CL*) se $\exists\, n \in \mathbb{N}$ tale che

$$\frac{1}{x^n} g(x) \quad \text{sia limitata in un intorno di } \pm\infty$$

Osservazione f supporto compatto $\Longrightarrow f$ decrescenza rapida $\Longrightarrow f$ crescenza lenta

Osservazione f_1, f_2 funzioni DR $\Longrightarrow f_1 \cdot f_2$ funzione DR.

Esempio 7.8.1 La funzione e^{-x^2} è a decrescenza rapida, anche se non ha supporto compatto.

Proposizione 7.8.1 *f funzione DR e g funzione CL $\Longrightarrow f \cdot g$ funzione DR*

Dimostrazione. Visto che g è funzione a crescenza lenta $\exists\, N \in \mathbb{N}$ per cui $x^{-N}g(x)$ è limitata in $\{|x| > M\}$, ovvero $|x^{-N}g(x)| < K$.

Allora per $f \cdot g$ vale

$$|x^n f(x)g(x)| = |x^{n+N}x^{-N}g(x)f(x)| < Kx^{n+N}|f(x)| \xrightarrow{|x|\to\infty} 0$$

in cui il limite è ottenuto proprio perché f è a decrescenza rapida.

□

Definizione 7.8.3 $\mathcal{S}(\mathbb{R}) := \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid \forall n \in \mathbb{N} \quad f^{(n)}$ è a decrescenza rapida $\}$ è chiamato *Spazio di Schwartz.*

Teorema 7.8.2 *Se $f \in \mathcal{S}(\mathbb{R})$, allora $\widehat{f} \in \mathcal{S}(\mathbb{R})$. Inoltre valgono le seguenti:*

- *Se $f \in \mathcal{S}(\mathbb{R})$, allora $(f^\wedge)^\vee = (f^\vee)^\wedge = f$.*
- *Se $f, g \in \mathcal{S}(\mathbb{R})$, allora vale*

$$\int_{\mathbb{R}} \widehat{f}(t)\widehat{g}(t)\, dt = \int_{\mathbb{R}} f(t)g(t)\, dt$$

- *$f : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R})$ è una isometria per la norma L^2.*

Osservazione

1. $\mathcal{S}$ è denso in L^2.
2. Se $f \in L^2$ e $\varphi_n \in \mathcal{S}$, allora $||\varphi_n - f||_{L^2} \longrightarrow 0$.
3. $\widehat{\varphi_n}$ è una successione di Cauchy in $L^2 \Longrightarrow \exists\, g \in L^2$ tale che $||\widehat{\varphi_n} - g|| \longrightarrow 0$ $(g = \widehat{f})$.

Definizione 7.8.4 Si dice *distribuzione temperata* un qualunque funzionale $T : \mathcal{S} \longrightarrow \mathbb{R}$ lineare continuo.

In questo caso il funzionale T è *continuo* se, presa una successione $\varphi_n \in \mathcal{S}$ infinitesima[2], vale $T(\varphi_n) \longrightarrow 0$, ovvero anche $T(\varphi_n)$ è infinitesima.

Definizione 7.8.5 Indichiamo con $\mathcal{S}'(\mathbb{R}) := \{T : \mathcal{S} \longrightarrow \mathbb{R} \mid T \text{ lineare continuo } \}$ lo spazio delle distribuzioni temperate.

Osservazione

1. $\mathcal{D}(\mathbb{R}) \subsetneq \mathcal{S}(\mathbb{R})$ infatti $e^{-x^2} \in \mathcal{S}(\mathbb{R}) \setminus \mathcal{D}(\mathbb{R})$.
2. Se $\{\varphi_n\} \in \mathcal{D}(\mathbb{R})$ e $\varphi_n \longrightarrow 0$ in $\mathcal{D}(\mathbb{R})$, allora $\varphi_n \longrightarrow 0$ in $\mathcal{S}(\mathbb{R})$.
3. Se $T \in \mathcal{S}'(\mathbb{R}) \Longrightarrow T_{\mathcal{D}} : \mathcal{D} \longrightarrow \mathbb{R}$ è lineare ($T : \mathcal{S} \longrightarrow \mathbb{R}$ lineare continuo).

Nota: Possiamo fare $\mathcal{D}(\Omega)$, con $\Omega \subseteq \mathbb{R}$, ma non $\mathcal{S}(\Omega)$, poiché ci serve sempre la condizione all'infinito, e nemmeno $\mathcal{S}'(\Omega)$. Quindi da ora in poi scriveremo solo $\mathcal{S}, \mathcal{S}'$, omettendo l'argomento "$\mathbb{R}$", che si ritiene sempre sottinteso.

Esempio 7.8.2 (Distribuzione temperata di tipo funzione) Se g è a crescenza lenta, $T_g : \mathcal{S} \longrightarrow \mathbb{R}$, definita come $< T_g, \varphi > = \int g\varphi$, definisce una distribuzione temperata. Dimostriamolo:

Se $\varphi \in \mathcal{S}(\mathbb{R})$, allora $g\varphi$ è sommabile, quindi $< T_g, \varphi >$ è ben definita.

La linearità è ovvia rispetto alla funzione test, e la continuità si verifica direttamente: se $\varphi_n \longrightarrow 0$ in $\mathcal{S}(\mathbb{R})$, allora

$$| < T_g, \varphi_n > | = \left| \int g(x)\varphi_n(x)\, dx \right| \leq \int |g(x)||\varphi_n(x)|\, dx =$$

visto che g è a crescenza limitata

$$= \int \frac{1}{1+x^2} \frac{g(x)}{1+|x|^s} (1+|x|^s)|\varphi_n(x)|(1+x^2)\, dx \leq$$

$$\leq \underbrace{\sup(1+|x|^s)(1+x^2)|\varphi_n(x)|}_{\xrightarrow[n\to\infty]{} 0} \cdot \int \frac{c}{1+x^2}\, dx$$

Quindi abbiamo appena dimostrato che

$$\lim_{n\to\infty} < T_g, \varphi >= 0$$

29/04/2014

Proposizione 7.8.2 *Siano $S, T \in \mathcal{S}'$, $c \in \mathbb{R}$, g funzione “abbastanza buona”*[3]. *Allora valgono le seguenti:*

1. $< S + T, \varphi > := < S, \varphi > + < T, \varphi >$
2. $< cT, \varphi > := \ c < T, \varphi >$
3. $< T', \varphi > := \ - < T, \varphi' >$
4. $< T(cx), \varphi(x) > := \ \frac{1}{|c|} < T(x), \varphi\left(\frac{x}{c}\right) >$
5. $< T(x - c), \varphi(x) > := < T(x), \varphi(x + c) >$
6. $< gT, \varphi > := < T, g\varphi >$

Dimostrazione. Molte di queste proprietà valgono per gli stessi motivi per cui valevano quando T ed S erano prese in $\mathcal{D}(\mathbb{R})$, quindi non le ripetiamo.

Poniamo l'accento su un paio di queste, per vedere cosa cambia:

3) questo caso è analogo al caso in cui si suppone φ a supporto compatto, se non ché l'integrale si annulla per un altro motivo:

$$< T_{f'}, \varphi >= \int_{\mathbb{R}} f'\varphi \, dx = \left[f\varphi\right]_{-\infty}^{+\infty} - \int_{\mathbb{R}} f\varphi' \, dx \quad \circledast$$

dove $|f(x)| \leq (1 + |x|^N) \cdot c$;

Poiché $\varphi \in \mathcal{S}$, vale

$$\lim_{x \to +\infty} (1 + |x|^N)\varphi(x) = 0$$

perciò il primo addendo in $\circledast$ è nullo, dunque rimane proprio quello che volevamo:

$$< T_{f'}, \varphi > = \ - < T_f, \varphi' >$$

6) anche questa definizione ha senso: se g è “abbastanza buona”, e $\varphi \in \mathcal{S}$, allora $g\varphi \in \mathcal{S}$. Per dimostrare quest'ultimo asserto dobbiamo mostrare che $g\varphi$ è a decrescenza rapida.

In effetti questo si può fare se ci ricordiamo cosa vuol dire *funzione a crescenza lenta*:

$$|x^n g(x)\varphi(x)| \leq |x|^n(1 + |x|^m) \cdot c|\varphi(x)| \xrightarrow{|x| \to +\infty} 0$$

[2]Nello spazio $\mathcal{S}$ una successione è infinitesima, ovvero $\varphi_n \longrightarrow 0$, se vale che $\forall k, \alpha \in \mathbb{N}$ $\sup_{\mathbb{R}}(1 + |x|^k)|\varphi^{(\alpha)}(x)| \longrightarrow 0$.

[3]In questo caso con il termine “abbastanza buona” richiediamo semplicemente che $T \in \mathcal{S} \Longrightarrow gT \in \mathcal{D}$.

mentre per le derivate:

$$|(g\varphi)'| = |g'\varphi + g\varphi'| \leq \left| x^n \Big(\underbrace{|g'(x)|\varphi(x)}_{\leq c_1(1+|x|^{m_1})} + \underbrace{|g(x)|\varphi'(x)}_{\leq c_2(1+|x|^{m_2})} \Big) \right|$$

e si conclude in maniera analoga a prima.

$\square$

Definizione 7.8.6 Definiamo le regole per le distribuzioni *coniugate*:

- $< \overline{T}, \varphi > := \overline{< T, \overline{\varphi} >}$
- $< \widehat{T}, \varphi > := < T, \widehat{\varphi} >$
- $< \overset{\vee}{T}, \varphi > := < T, \overset{\vee}{\varphi} >$

Abbiamo detto che se $\varphi \in \mathcal{S}$, allora $\widehat{\varphi} \in \mathcal{S}$ e $(\widehat{\varphi})^\vee = \varphi$ (anche il viceversa vale).
Anche in $\mathcal{S}'$ vale $(\widehat{T})^\vee = T$, infatti per le definizioni appena date abbiamo

$$< (\widehat{T})^\vee, \varphi > = < \widehat{T}, \overset{\vee}{\varphi} > = < T, (\widehat{\varphi})^\vee > = < T, \varphi >$$

Definizione 7.8.7 Diciamo che in $\mathcal{S}'$ vale $T_n \xrightarrow{n \to +\infty} T$ se e solo se

$$\forall \varphi \in \mathcal{S} \qquad < T_n, \varphi > \xrightarrow{n \to +\infty} < T, \varphi >$$

Osservazione Vale anche la seguente proprietà: $T_n \longrightarrow T \quad \Longrightarrow \quad \widehat{T_n} \longrightarrow \widehat{T}$.

Esempio 7.8.3 Ci ricordiamo la δ? Questa è una distribuzione temperata (non lo dimostriamo ma è facile da vedere), allora chi è la $\widehat{\delta}$?

$$\widehat{\varphi}(t) = \int_{\mathbb{R}} e^{-2\pi i t x} \varphi(x)\, dx$$

$$\Longrightarrow \quad < \widehat{\delta}, \varphi > = < \delta, \widehat{\varphi} > = \widehat{\varphi}(0) = \int_{\mathbb{R}} \varphi(x)\, dx = < 1, \varphi >$$

quindi $\widehat{\delta} = 1$!

Allo stesso modo si può dimostrare che $\overset{\vee}{\delta} = 1$ e che $\widehat{1} = \overset{\vee}{1} = \delta$.

Teorema 7.8.3 *Se $T \in \mathcal{S}'$, allora*

1. $xT(x) = 0 \quad \Longleftrightarrow \quad \exists\, c \in \mathbb{C}$ *tale che* $T = c \cdot \delta$
2. *se* $T' = 0$, *allora* $\exists\, c \in \mathbb{C}$ *tale che* $T = T_c$.

Dimostrazione. Dimostriamo i due punti separatamente:

1. $\boxed{\Longleftarrow}$ Sia $T = c\delta$. Allora

$$< xT(x), \varphi(x) > = < T(x), x\varphi(x) > = \ c \cdot 0 \cdot \varphi(0) = 0$$

$\boxed{\Longrightarrow}$ Sia $g \in \mathcal{S}$, con $g(0) = 0$. Allora si dimostra facilmente che

$$\frac{g(x)}{x} \in \mathcal{C}^\infty \qquad\qquad \frac{g(x)}{x} \in \mathcal{S}'$$

d'altra parte dividere per x ci fa solo bene per stare in $\mathcal{S}'$.

Sia $\psi \in \mathcal{S}$, con $\psi(0) = 1$. Se $g(0) = 0$ allora $< T, g >= 0$ perché

$$< T(x), g(x) > = < T(x), x\frac{g(x)}{x} > = < xT(x), \frac{g(x)}{x} > = \ 0$$

Ora, sia $\varphi \in \mathcal{S}$ qualsiasi.

$$g(x) = \varphi(x) - \varphi(0)\psi(x) \quad \Longrightarrow \quad g(0) = 0$$

dunque per quanto detto sopra abbiamo $< T, g >= 0$; inoltre vale

$$0 =< T, g > = < T(x), \varphi(x) - \varphi(0)\psi(x) > = < T(x), \varphi(x) > -\varphi(0) < T(x), \psi(x) >$$

$$\Longrightarrow < T, \varphi >= \varphi(0) \cdot \underbrace{< T, \psi >}_{:=c} = c < \delta, \varphi >$$

come volevamo.

2. Supponiamo $T' = 0$.

Sicuramente vale sempre[4]

$$\widehat{T'}(t) = 2\pi i t \widehat{T}(t)$$

Se $T' = 0$, anche $\widehat{T'} = 0$, quindi $t\widehat{T}(t) = 0$.

Per il punto (1) avremo quindi $t\widehat{T}(t) = c\delta$, da cui

$$T = (\widehat{T})^\vee = (c\delta)^\vee = c$$

[4] Avevamo visto che valeva per le funzioni ordinarie; vale anche per le distribuzioni temperate, dato che scarichiamo tutto sulla funzione φ.

□

Definizione 7.8.8 $T \in \mathcal{S}'$ si dice[5] *nulla* nell'intervallo (a,b) se $\forall \varphi \in \mathcal{S}$, con $supp\varphi \subseteq (a,b)$ vale $< T, \varphi > = 0$.

Teorema 7.8.4 *Se $T \in \mathcal{D}'(\mathbb{R})$ è tale che[6] $T' = 0$ in $]a,b[$ (nel senso della precedente definizione), allora T è costante in $]a,b[$, ovvero $\exists\, c \in \mathbb{R}$ per cui $T - cT_1 = 0$ in $]a,b[$.*

Dimostrazione. Sia $g \in \mathcal{D}(\mathbb{R})$, con $supp\, g \subseteq]a,b[$. Supponiamo $\int_a^b g = 0$.

Prendiamo $\psi \in \mathcal{D}(\mathbb{R})$, con $supp\, \psi \subseteq]a,b[$ e $\int_a^b \psi = 1$.

Definiamo anche

$$\theta(x) := \int_a^x g(t)\; dt$$

che risulta quindi $\theta \in \mathcal{D}(\mathbb{R})$ con $supp\, \theta \subseteq]a,b[$.

Sia ora $\varphi \in \mathcal{D}(\mathbb{R})$ qualsiasi, con $supp\, \varphi \subseteq]a,b[$. Vogliamo mostrare che per qualche c vale

$$< T, \varphi >= c\int_a^b \varphi$$

Sia $K = \int_a^b \varphi$. Allora

$$\int_a^b (\underbrace{\varphi - K\psi}_{g}) = 0$$

Ricordando la definizione di θ, da cui $\theta' = g$, ricaviamo infine

$$0 =< T', g >= - < T, \theta' >= - < T, g >$$

$$\Longrightarrow\; 0 =< T, g >=< T, \varphi - K\psi >=< T, \varphi > -K < T, \psi >=< T, \varphi > - \int_a^b \varphi \cdot c$$

$$\Longrightarrow < T, \varphi >= c\int_a^b \varphi$$

□

[5] Questa definizione si può dare analoga per le distribuzioni in $\mathcal{D}$.
[6] Ricordiamo che $\mathcal{S}' \subseteq \mathcal{D}'$, quindi T può anche essere presa in $\mathcal{S}'$.

Esempio 7.8.4 La trasformata di Fourier di $f(x) = \frac{1}{1+x^2}$.[7]

$$\widehat{f}(t) = \int_{-\infty}^{+\infty} e^{-2\pi i t x}\frac{1}{1+x^2}\,dx =$$

$$= \int_{-\infty}^{+\infty} \frac{\cos(2\pi t x)}{1+x^2}\,dx + i\int_{-\infty}^{+\infty} \frac{\sin(2\pi t x)}{1+x^2}\,dx = \int_{-\infty}^{+\infty} \frac{\cos(2\pi t x)}{1+x^2}\,dx$$

in cui l'ultima uguaglianza è ottenuta perché la funzione seno è dispari.

L'integrale rimanente si potrebbe calcolare con il metodo dei residui, ma noi faremo in un altro modo: ricordando che $(xf(x))^\vee(t) = \frac{-1}{2\pi i}\widehat{f}'(t)$, ricaviamo

$$\big((1+x^2)f(x)\big)^\vee = \widehat{1} = \delta \quad\Longrightarrow\quad \delta = \widehat{f}(t) - \frac{1}{4\pi^2}\widehat{f}''(t)$$

$$\Longrightarrow\quad \widehat{f}'' - 4\pi^2\widehat{f} = -4\pi^2\delta$$

ma $-4\pi^2\delta = 0$ in $]-\infty,0[\cup]0,+\infty[$, da cui

$$\widehat{f}'' = -4\pi^2\widehat{f}$$

Attenzione: stiamo considerando un'equazione differenziale definita du un insieme non connesso, con le conseguenze che ne derivano su come è fatta la soluzione.

Cerchiamo dunque $\widehat{f}$ nella forma

$$\widehat{f}(t) = Ae^{-2\pi t} + Be^{2\pi t}$$

perché deve essere $\lambda^2 - 4\pi^2 = 0$, ovvero $\lambda = \pm 2\pi$. Dunque

$$\widehat{f}(t) = \begin{cases} Ae^{-2\pi t} & \text{se } t < 0 \\ Be^{2\pi t} & \text{se } t > 0 \end{cases}$$

Ora, in realtà f è pari, dunque anche $\widehat{f}$ è pari, quindi dobbiamo prendere $A = B$, per cui la nostra funzione è

$$\widehat{f}(t) = Ae^{-2\pi|t|}$$

Quindi vale anche $\widehat{f}(0) = A$, dove

$$\widehat{f}(0) = \int_{\mathbb{R}} \frac{1}{1+x^2}\,dx = \pi \quad\Longrightarrow\quad \widehat{f}(t) = \pi e^{-2\pi|t|}$$

[7] Prima abbiamo ottenuto la trasformata di Fourier di 1, che **non** è una funzione sommabile quindi abbiamo usato la strada che abbiamo usato. Stavolta f è sommabile, quindi possiamo fare le trasformate in maniera tradizionale.

Posto $\widehat{f}(t) = g(t)$, possiamo anche verificare che $\check{g}(x) = \frac{1}{1+x^2}$, infatti

$$\check{g}(x) = \int_{\mathbb{R}} e^{2\pi itx} \cdot \pi e^{-2\pi|t|}\, dt = \underbrace{\int_{-\infty}^{0} e^{2\pi itx} \cdot \pi e^{-2\pi|t|}\, dt}_{\circledast} + \underbrace{\int_{0}^{+\infty} e^{2\pi itx} \cdot \pi e^{-2\pi|t|}\, dt}_{\circledcirc}$$

per sapere quanto vale $\circledast$ calcoliamo il valore di $\circledcirc$:

$$\circledcirc = \int_0^{+\infty} \pi e^{2\pi(ix-1)t}\, dt = \frac{\pi}{2\pi(ix-1)}\Big[e^{2\pi(ix-1)t}\Big]_0^{+\infty} = \frac{1}{2(1-ix)}$$

Visto che l'integrale su tutto $\mathbb{R}$ deve valere $[2(1+ix)]^{-1}$, $\circledast$ deve valere $[1+x^2]^{-1}$, che è proprio quello che volevamo dimostrare.

05/05/2014

Esempio 7.8.5 La trasformata di Fourier di

$$H_d(x) = \begin{cases} \frac{1}{2} & \text{se } x > 0 \\ -\frac{1}{2} & \text{se } x < 0 \end{cases}$$

Sappiamo già che la derivata della funzione $H_d(x)$ è la funzione δ.
Sappiamo che $(T')^\wedge(t) = 2\pi it\widehat{T}(t)$, quindi nel nostro caso

$$1 = \widehat{\delta}(t) = (H')^\wedge(t) = 2\pi it\widehat{H}(t)$$

$$\Longrightarrow \quad \widehat{H}(t) = \frac{1}{2\pi it}$$

Ma in realtà questa ultima divisione non è del tutto lecita, visto che l'uguaglianza scritta sopra è una uguaglianza tra distribuzioni, non tra valori reali! Quindi ricaviamo l'ultima relazione scritta, che è giusta, in modo più corretto.

Sicuramente $t \cdot \frac{1}{t} = 1$, quindi possiamo scrivere

$$2\pi it\widehat{H}(t) = 1 \quad \Longrightarrow \quad 2\pi it\widehat{H}(t) - t\frac{1}{t} = 0$$

$$\Longrightarrow \quad t\left(2\pi i\widehat{H}(t) - \frac{1}{t}\right) = 0$$

Abbiamo già dimostrato che se T è una distribuzione vale $tT(t) = 0 \Longleftrightarrow T(t) = a\delta$. Applicando questo risultato al nostro caso abbiamo

$$2\pi i\widehat{H}(t) - \frac{1}{t} = a\delta \quad \Longrightarrow \quad \widehat{H}(t) = \frac{1}{2\pi it} + b\delta$$

dove $b = \frac{a}{2\pi i}$.

Ricordiamo che se $\varphi(x)$ è una funzione reale dispari otteniamo

$$\widehat{\varphi}(t) = \int_{-\infty}^{+\infty} e^{-2\pi itx}\varphi(x)\, dx = \int_{-\infty}^{+\infty} i\sin(2\pi tx)\varphi(x)\, dx$$

il che ci dice che la trasformata di Fourier di una funzione dispari è una funzione a valori immaginari puri!

In effetti la nostra H è reale dispari[8], quindi $\widehat{H}$ è immaginaria pura sulle funzioni φ reali (cioè tali che $< H, \varphi > \in \mathbb{R}$), vediamo perché:

- H dispari, g pari $\Longrightarrow < H, g >= 0$, infatti

$$< H(x), g(x) >=< H(x), g(-x) >=< H(-x), g(x) >= - < H(x), g(x) >$$

- H dispari $\Longrightarrow$ $\widehat{H}$ dispari, infatti

$$< \widehat{H}(-x), \varphi(x) >=< H(-x), \widehat{\varphi}(x) >= - < H(x), \widehat{\varphi}(x) >= - < \widehat{H}(x), \varphi(x) >$$

- H reale dispari $\Longrightarrow$ $\widehat{H}$ è dispari, anche se questo è un po' più articolato da vedere. Trasformiamo la funzione test φ come somma di funzioni pari e dispari

$$\varphi = \frac{1}{2}\underbrace{\big(\varphi(x) + \varphi(-x)\big)}_{\text{pari}} + \frac{1}{2}\underbrace{\big(\varphi(x) - \varphi(-x)\big)}_{\text{dispari}}$$

  da cui possiamo usare i punti precedenti ed ottenere

$$< \widehat{H}, \varphi >= \frac{1}{2} < \widehat{H}(x), \varphi(x) - \varphi(-x) >= \frac{1}{2} < H(x), \widehat{\varphi}(x) - \widehat{\varphi}(-x) >$$

  sapendo che $\widehat{\varphi}(x) - \widehat{\varphi}(-x)$ è immaginaria pura (perché trasformata di una funzione dispari) la scriviamo come $if(x)$, quindi

$$\frac{1}{2} < H(x), \widehat{\varphi}(x) - \widehat{\varphi}(-x) >= \frac{1}{2} < H(x), if(x) >= \frac{1}{2}i < H(x), f(x) >= \frac{1}{2}ib$$

$$b \in \mathbb{R}$$

Dunque cosa abbiamo: H è dispari, quindi $\widehat{H}$ risulta dispari con valori immaginari puri. Ma δ è pari, quindi per avere l'uguaglianza di sopra b deve essere nullo, ovvero a deve essere nullo, quindi

$$\Longrightarrow \quad \widehat{H}(t) = \frac{1}{2\pi it}$$

come volevamo dimostrare.

[8] Ricordiamo che in questo caso H dispari ci dice che $H(x) = H(-x)$ come distribuzione, non come funzione.

Esempio 7.8.6 Calcoliamo le trasformate di $|x| \cdot x^{n-1}$ e $\frac{1}{x^n}$.

Lemma *Per ogni distribuzione T vale la seguente uguaglianza:*

$$(x^n T(x))^\wedge(t) = \left(\frac{-1}{2\pi i}\right)^n \left(\widehat{T}\right)^{(n)}(t)$$

Usiamo il Lemma:

$$(x^n H(x))^\wedge(t) = \left(\frac{-1}{2\pi i}\right)^n \frac{1}{2\pi i}\left(\frac{d}{dt}\right)^n \left(\frac{1}{t}\right)$$

La derivata n-esima di t^{-1} risulta essere $(-1)^n n! t^{-n-1}$, dimostrabile facilmente per induzione. Quindi risulta

$$(x^n H(x))^\wedge(t) = \left(\frac{1}{2\pi i}\right)^n \frac{n!}{2\pi i}\frac{1}{t^{n+1}}$$

In particolare, per $n = 1$ abbiamo

$$(xH(x))^\wedge(t) = \frac{-1}{4\pi^2}\frac{1}{t^2}$$

da cui possiamo ricavare

$$\left(\frac{1}{2}|x|\right)^\wedge(t) = \frac{-1}{4\pi^2}\frac{1}{t^2}$$

Nel caso della nostra H abbiamo che

$$|x|x^{n-1} = 2x^n H(x)$$

da cui

$$\left(|x|x^{n-1}\right)^\wedge(t) = 2\left(x^n H(x)\right)^\wedge(t) = \frac{2n!}{(2\pi i)^{n+1}}\frac{1}{t^{n+1}}$$

se sostituiamo n con $n-1$ risulta

$$\frac{1}{t^n} = \frac{(2\pi i)^n}{2(n-1)!}\left(|x|x^{n-2}\right)^\wedge(t) := \widehat{S}(t)$$

Possiamo calcolare l'antitrasformata di $\widehat{S}(t)$:

$$\left(\frac{(-1)^n}{t^n}\right)^\wedge(x) = \frac{(2\pi i)^n}{2(n-1)!}|x|x^{n-2}$$

per $n = 1$ risulta

$$\left(\frac{(-1)}{t}\right)^\wedge(x) = \frac{2\pi i}{2} sgn(x) = 2\pi i H(x)$$

il che conferma anche i calcoli fatti in precedenza; Abbiamo fatto una specie di “verifica”.

Capitolo 8

Ricerca Operativa - Parte 3

05/05/2014

Tra due variabili aleatorie X e Y, che esprimono valori monetari, avevamo deciso di definire X *preferibile* a Y se $\mathbb{E}(X) > \mathbb{E}(Y)$, oppure *indifferente* (indicato con $X \sim Y$) se $\mathbb{E}(X) = \mathbb{E}(Y)$.

Esempio 8.0.7 (Assicurazioni) Pagando 100 euro si gioca ad un gioco con le seguenti regole: con probabilità $\frac{1}{10000}$ vinciamo 20000 euro, nel resto dei casi vinciamo 0. Conviene giocare?

Se X indica l'evento "gioco", il valore atteso è $\mathbb{E}(X) - 100 = 2 - 100 = -98$, se Y indica l'evento "non gioco", il valore atteso è $\mathbb{E}(Y) = 0$.

Visto così ci conviene di molto <u>non</u> giocare.

Questo però non è così detto nella vita reale, visto che le assicurazioni funzionano esattamente così: paghiamo una cifra accettabile ogni anno per prevenire un evento con probabilità molto trascurabile.

8.1 Paradosso di San Pietroburgo

Nel 1738 Nicolas Bernoulli propose al cugino Daniel il **Paradosso di San Pietroburgo**: un tizio lancia una moneta finché non esce per la prima volta testa. Se la prima testa esce al lancio n-esimo ricevo un premio di 2^{n-1}.

Qual'è il prezzo equo da pagare per questo gioco? Solitamente ragioniamo in termini di speranza della vincita, calcoliamola:

La probabilità che la prima testa esca al lancio n-esimo è $\left(\frac{1}{2}\right)^n$, quindi la speranza risulta

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 2^{n-1} = \sum_{n=1}^{\infty} \frac{1}{2} = +\infty$$

Quindi il prezzo “equo” può essere alto a piacere! Potremmo chiedere anche 10000000 euro di prezzo; il problema è che sarà difficile vincere più di qualche euro, visto che la probabilità di vincere grandi quantità di denaro è molto piccola (esattamente inversamente proporzionale).

06/05/2014

Vediamo quanto alti potrebbero essere i guadagni, per farlo calcoliamo la probabilità che la testa esca prima del quinti tiro (che sembra ragionevole):

$$\mathbb{P}(X \leq 5) = 1 - \left(\frac{1}{2}\right)^5 = 1 - \frac{1}{32} = \frac{31}{32} \simeq 0.97$$

quindi con una probabilità molto alta (circa il 97%) non vinceremo più di pochi euro, a fronte di una speranza matematica infinita!

Diamo un nome alle variabili in gioco per fare uno studio più approfondito: $m =$ patrimonio del giocatore; $f =$ tariffa richiesta per giocare.

Osservazione Cramer, tra i suoi molti studi, individuò la regola $u(x + \Delta x) = k\frac{\Delta x}{x}$, dove $u(x)$ indica il nostro patrimonio. Essa esprime il fatto che la variazione di patrimonio è lineare dispetto alla variazione di denaro, ovvero se abbiamo un patrimonio molto alto, perché una variazione di esso sia sensibile dobbiamo vincere/perdere molto denaro.

Da questa formula ricaviamo infatti

$$u'(x) = k\frac{1}{x} \quad \Longrightarrow \quad u(x) = k \log x + c \simeq \log x$$

L'idea di per sé non è nulla di che, lo sapevamo già tutti, ma la genialità di Cramer è stata quella di riuscire a porlo in formule matematiche!

Il patrimonio che ci troveremo in tasca dopo aver giocato una partita è calcolabile come $m - f + 2^n$. Utilizzando l'idea di Cramer, supponiamo $u(x) = \log x$; possiamo calcolare la speranza di vincita scegliendo di giocare (X esprime la scelta “gioco”) in modo diverso da prima:

$$\mathbb{E}(u(X)) = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \log\left(m - f + 2^n\right)$$

anche se la somma della serie non è calcolabile esplicitamente, è facile verificare che la serie in effetti converge.

Se vogliamo calcolare la speranza di vincita scegliendo di non giocare (Y esprime la scelta “non gioco”) è anche più semplice:

$$\mathbb{E}(u(Y)) = u(Y) = \log m$$

Abbiamo detto che il gioco è equo quando questi due valori sono uguali (ovvero quando è equivalente giocare o non giocare). Grazie a questo ragionamento possiamo calcolare il valore "equo" della tariffa richiesta per giocare f, che evidentemente dipende dal patrimonio del singolo giocatore, ed è proprio ciò che fa Bernoulli, calcolando alcuni valori:

m	200	1000	10^5	10^9
f	8.72	10.95	17.56	24.20

Approfondimento: Cosa accadrebbe se la vincita del gioco fosse 2^{2^n} invece di 2^{n-1} (mantenendo la modalità di vincita invariata)?

Sicuramente variano i valori calcolati delle speranze: nel caso si decida di giocare otteniamo

$$\mathbb{E}(u(X)) = \sum_{n=1}^{\infty} \frac{1}{2^n} \log\left(m - f + 2^{2^n}\right)$$

Quindi per scoprire il valore equo della scommessa iniziale dovremmo rifare i conti con queste nuove funzioni, ottenendo valori molto diversi.

Esercizio 8.1.1 Provare a calcolare il valore equo di entrata nel caso di vincita 2^{2^n}.

8.2 Funzione utilità

La "funzione utilità", che prima abbiamo indicato con $u(x)$, deve essere crescente e concava, difatti il logaritmo lo è, ma cosa succedere se u è limitata? e abbiamo $|u(x)| \leq M$ allora

$$\mathbb{E}(u(X)) = \sum p_k u(x_k) \leq M \sum p_k = M$$

anche la speranza matematica risulta limitata.

Osservazione (Attitudine al rischio) Se prendiamo

$$u(x) = 1 - e^{-\frac{x}{a}}$$

essa ha un asintoto orizzontale in 1, quindi è limitata.

Cosa esprime il valore a? Se usiamo Taylor

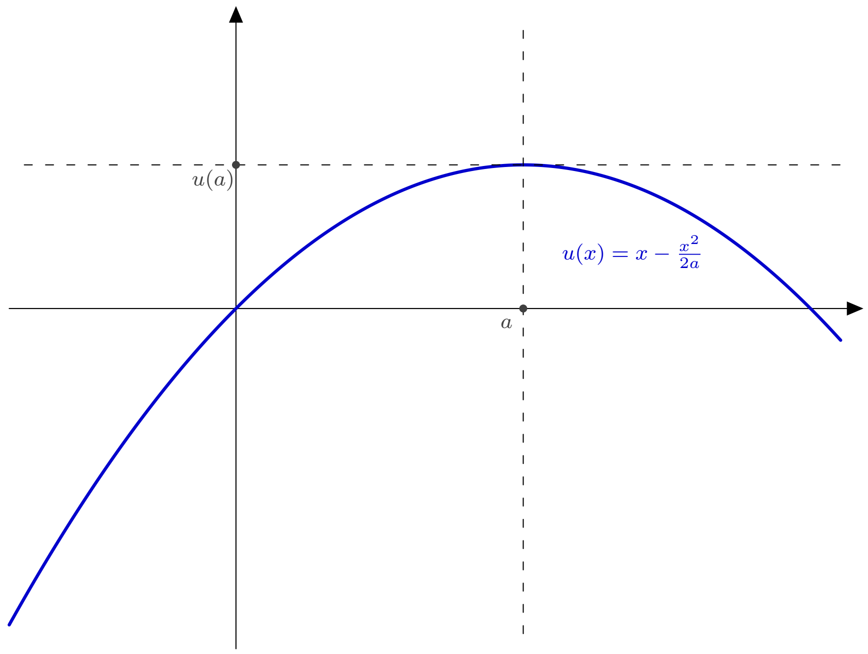
$$e^{-\frac{x}{a}} = 1 - \frac{x}{a} + \frac{1}{2}\frac{x^2}{a^2} + o(x^2)$$

$$\Longrightarrow \quad u(x) = \frac{1}{a}\left(x - \frac{1}{2a}x^2 + o(x^2)\right)$$

quindi il termine di secondo grado in $u(x)$ è tanto meno rilevante quanto più è grande a, quindi questo a è in qualche modo una "misura dell'attitudine al rischio".

Esempio 8.2.1 (Utilità quadratica) Poniamo $u(x) = x - \frac{x^2}{2a}$.

Purtroppo questa funzione è concava ma non crescente, il grafico infatti è di questo tipo



Quindi una funzione utilità di questo tipo può essere usata se stiamo attenti che i valori monetari in gioco non superino mai il valore a.

Esempio 8.2.2 Rappresentiamo un gioco in forma tabulare: le colonne d_i rappresentano le decisioni possibili (d_5=non giocare), le righe rappresentano gli eventi possibili E_1,E_2, di cui riportiamo le probabilità nell'ultima colonna, le ultime righe rappresentano i valori attesi decidendo di giocare.

La funzione utilità usata è $u_i(x) = x - \frac{x^2}{2a_i}$, con attitudini al rischio dei giocatori

$a_1 = 25000$, $a_2 = 250000$.

	d_1	d_2	d_3	d_4	d_5	$\mathbb{P}$
E_1	30	-960	10060	10000	0	$\frac{1}{2}$
E_2	-10	1000	-10000	-9980	0	$\frac{1}{2}$
$\mathbb{E}(G)$	10	20	30	10	0	
$\mathbb{E}(u_1(G))$	9.99	0.784	-1982	-5496	0	
$\mathbb{E}(u_2(G))$	9.999	18	-171	-4590	0	

Qualunque sia u il “non giocare” risulta preferibile ad un gioco equo, vediamo perché:

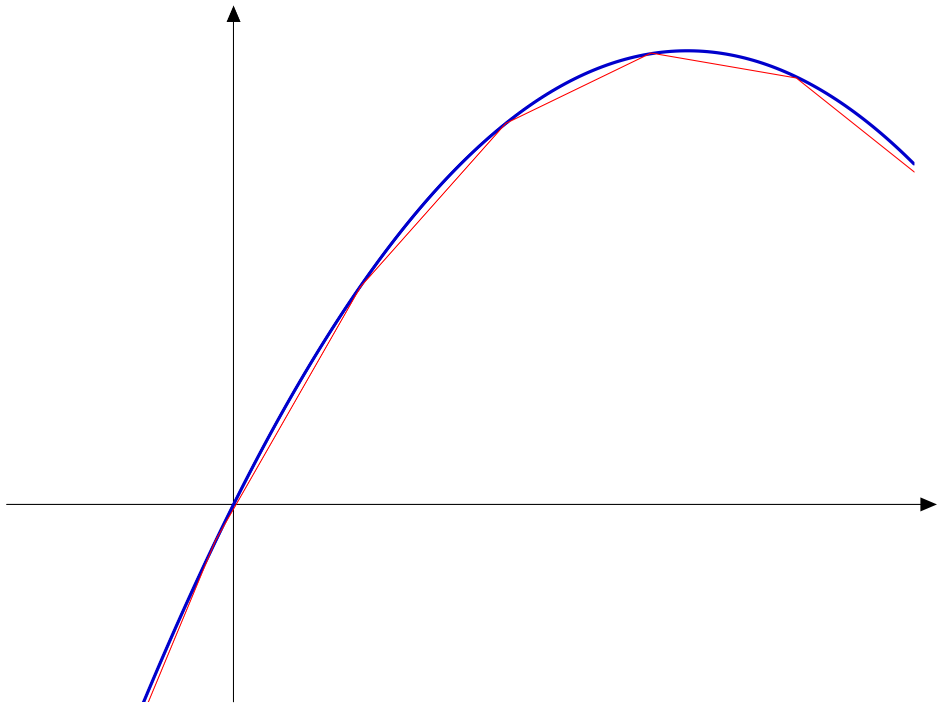
Chiamiamo X_1 il guadagno netto se non giochiamo (che risulta nullo) e X_2 il guadagno netto se giochiamo.

Chiamiamo poi gli eventi possibili y e z, che accadono con probabilità p e $1 - p$ rispettivamente. Nel gioco equo risulta $py + (1 - p)z = 0$.

Visto che u è funzione concava vale sicuramente che

$$u(py + (1 - p)z) > pu(y) + (1 - p)u(z)$$

proprio perché u concava significa che le “corde” stanno al di sotto di essa. Si vede bene con un grafico:



Osservazione Proviamo ad analizzare l'utilità di un guadagno aleatorio X, in funzione della speranza $m = \mathbb{E}(X)$ e della varianza $\sigma^2 = var(X)$.

Usiamo ancora la funzione $u(x) = x - \frac{x^2}{2a}$. Il valore atteso risulta

$$\mathbb{E}(u(X)) = \int_{-\infty}^{a} \left(x - \frac{x^2}{2a}\right) \, dF(x)$$

in cui F= funzione di ripartizione di X. Calcoliamo questo valore atteso:

$$\mathbb{E}(u(X)) = \int_{-\infty}^{a} x \, dF(x) - \frac{1}{2a} \int_{-\infty}^{a} x^2 \, dF(x) = \mathbb{E}(X) - \frac{1}{2a}\mathbb{E}\left(X^2\right) =$$

$$= \mathbb{E}(X) - \frac{1}{2a}\mathbb{E}\left(X^2\right) + \frac{1}{2a}\big(\mathbb{E}(X)\big)^2 - \frac{1}{2a}\big(\mathbb{E}(X)\big)^2 =$$

$$= m - \frac{1}{2a}\sigma^2 - \frac{1}{2a}m^2 = u(m) - \frac{1}{2a}\sigma^2$$

Esempio 8.2.3 (Diversificazione di investimenti) Supponiamo di avere da investire un certo importo (20000 euro). Ci vengono proposte due lotterie indipendenti X, Y con stessa media m, stessa varianza σ^2, e stessa quota di partecipazione pari a metà del nostro capitale (10000 euro).

Possiamo decidere di giocare tutto nella prima (o seconda) lotteria, ottenendo un'utilità attesa (usando sempre la $u(x)$ di prima)

$$\mathbb{E}(u(2X)) = \mathbb{E}(u(2Y)) = u(2m) - \frac{1}{2a}2^2\sigma^2 = u(2m) - \frac{2}{a}\sigma^2$$

oppure decidere di giocare in entrambe le lotterie una volta sola, ottenendo un'utilità attesa

$$\mathbb{E}(u(X+Y)) = u(2m) - \frac{1}{2a}\left(1^2 + 1^2\right)\sigma^2 = u(2m) - \frac{1}{a}\sigma^2$$

che è decisamente meglio, perché si abbassa la varianza.

Osservazione Il discorso appena fatto per due lotterie vale anche per un numero qualsiasi di lotterie differenti, anzi il fenomeno si accentua sempre di più: conviene sempre partecipare a molte lotterie diverse piuttosto che puntare tutto su una!

Questo è quello che fanno le assicurazioni in effetti: stipulando molte polizze da pochi euro guadagnano molto di più che non stipulando poche polizze con importi molto alti.

12/05/2014

Analizziamo meglio questa osservazione: quando conviene alla compagnia accettare un assicurato?

r = importo del premio ricevuto.

S = importo liquidato in caso di evento (con probabilità p).

Avremo dunque un guadagno r con probabilità p, o un guadagno $r - S$ con probabilità $1 - p$. Possiamo quindi calcolare la speranza:

$$\mathbb{E}(X) = r - rp + rp - Sp = r - Sp$$

da cui $\mathbb{E}(X) > 0 \iff r > Sp$.

In effetti questo risultato era abbastanza prevedibile in questo caso, visto che si ritorna a studiare quando il prezzo assicurativo è equo.

Questo modello non tiene però conto del fatto che l'assicurazione deve poter guadagnare qualcosa. Per tenere conto di questo deve essere messa in gioco qualche questione di utilità.

Consideriamo la funzione $u(x) = x - \frac{x^2}{2a}$, come al solito, e introduciamo dei nuovi termini:

$r = pS$ lo chiamiamo *premio puro*.

$r = kS$, con $k > p$, lo chiamiamo *premio caricato*.

$\gamma = k - p$ lo chiamiamo *caricamento unitario*.

In questo caso i valori assicurativi diventano kS con probabilità $1 - p$, oppure $kS - S$ con probabilità p. Calcoliamo quindi la speranza:

$$\begin{aligned}
\mathbb{E}(u(X)) &= pu(kS - S) + (1-p)u(kS) = \\
&= p\left(kS - S - \frac{1}{2a}(kS - S)^2\right) + (1-p)\left(kS - \frac{1}{2a}(kS)^2\right) = \\
&= kS - pS - \frac{1}{2a}\left[k^2S^2 + pS^2 - 2pkS^2\right] = \\
&= kS - pS - \frac{1}{2a}S^2\left(k^2 - 2pk + p^2 + p - p^2\right) = \\
&= \gamma S - \frac{S^2}{2a}\left(\gamma^2 + p(1-p)\right)
\end{aligned}$$

Si desidera $\mathbb{E}(u(X)) > 0$, quindi

$$\mathbb{E}(u(X)) > 0 \iff \frac{S}{2a}(\gamma^2 + p(1-p)) < \gamma \iff$$

$$S < \frac{2a\gamma}{\gamma^2 + p(1-p)}$$

per capire come funziona facciamo un esempio.

Esempio 8.2.4 Poniamo che l'assicurazione chieda un pagamento di $k = 2p$, con $\gamma = p$. La soglia del premio per la convenienza dell'utilità diviene quindi calcolabile con la formula di sopra:

$$S < \frac{2a\gamma}{\gamma^2 + p(1-p)} = \frac{2ap}{p^2 + p(1-p)} = 2a$$

Aumentiamo ancora le quantità e poniamo $k = 3p$, stavolta con $\gamma = 2p$. La soglia stavolta diviene

$$S < \frac{2a\gamma}{\gamma^2 + p(1-p)} = \frac{4ap}{4p^2 + p - p^2)} = \frac{4a}{3p+1} \xrightarrow{p \simeq 0} 4a$$

In generale, visto che si assicurano eventi molto improbabili piuttosto che quelli molto probabili, gli ordini di grandezza delle quantità in gioco sono $p(1-p) = \sigma^2 \simeq p$. Quindi la soglia di rischio si può approssimare più semplicemente come

$$S < \frac{2a\gamma}{\gamma^2 + \sigma^2} \simeq \frac{2a\gamma}{\sigma^2} \simeq \frac{2a\gamma}{p}$$

Ovvero conviene accettare l'assicurato se

$$\gamma > \frac{\sigma^2 S}{2a} \left(\simeq \frac{pS}{2a} \right)$$

Nota: la quantità a, che entra in gioco in ogni calcolo fatto negli ultimi paragrafi, abbiamo già detto che indica l'attitudine al rischio, in qualche modo. Questo vuol dire che a è un parametro che dipende da moltissimi fattori: il capitale disponibile dell'azienda, il numero di persone assicurate, le quote in gioco in ogni assicurazione, la "bontà" dell'operatore che concede l'assicurazione (o del manager del settore), e molti altri ancora.

Passiamo al problema **duale**: Quando conviene sottoscrivere un'assicurazione? (vediamo tutto dal punto di vista dell'assicurato, e non dell'assicurazione)

Schematizziamo la situazione seguente: dobbiamo decidere se stipulare l'assicurazione della macchina (decisione 2) oppure no (decisione 1). Chiamiamo poi E l'evento "la macchina brucia e la devo ricomprare" (con probabilità p), e analizziamo cosa accade a seconda della decisione presa, in cui le quantità di denaro sono: S = costo della macchina nuova, kS = quota di assicurazione (in cui k è dell'ordine di p).

	d_1 (*no*)	d_2 (*si*)	$\mathbb{P}$
E	$-S$	$-kS$	p
$\overline{E}$	0	$-kS$	$1-p$

come al solito usiamo la funzione utilità

$$u(X) = x - \frac{x^2}{2a}$$

e introduciamo le due quantità g_1 = utilità attesa in caso di d_1, calcolata come

$$g_1 = pu(-S) + (1-p)u(0) = p\left(-S - \frac{S^2}{2a}\right) = -pS\left(1 + \frac{S}{2a}\right)$$

oppure g_2 = utilità attesa in caso di d_2, calcolata come

$$g_2 = pu(-kS) + (1-p)u(-kS) = u(-kS) = -kS - \frac{k^2S^2}{2a}$$

Confrontiamo le due quantità per decidere quale conviene:

$$g_1 - g_2 = (k-p)S - \frac{S^2}{2a}(p - k^2)$$

visto che $k \simeq p$, possiamo approssimare $p - k^2 \simeq p$ ed ottenere

$$g_1 - g_2 \simeq (k-p)S - \frac{pS^2}{2a} = \gamma S - \frac{pS^2}{2a}$$

Perché ci convenga g_2 (ovvero conviene assicurarci) dobbiamo avere $g_1 - g_2 < 0$, ovvero

$$g_1 - g_2 < 0 \Longleftrightarrow \gamma S - \frac{pS^2}{2a} < 0 \Longleftrightarrow \gamma < \frac{pS}{2a}$$

che risulta essere proprio il risultato "duale" di quello ottenuto facendo i conti da parte dell'assicurazione, invece che dell'assicurato.

Quando facevamo i conti dalla parte dell'assicurazione, l'attitudine al rischio a era chiaramente diversa da quella del singolo. Se chiamiamo a_1 l'attitudine al rischio della compagnia e a_2 quella del singolo individuo, è chiaro che si avrà $a_2 < a_1$, visto che l'assicurazione <u>deve</u> avere più attitudine al rischio (vive proprio su quello) rispetto ad un singolo che non vorrebbe rischiare la sua (magari unica) macchina.

Le soglie di rischio che abbiamo calcolato per entrambe le parti sono

$$\gamma_1 < \frac{pS}{2a_1} \qquad \gamma_2 < \frac{pS}{2a_2}$$

per cui sembrano le stesse. Ma tenendo conto della disuguaglianza delle parti, ovvero $a_2 < a_1$, troviamo $\gamma_1 < \gamma_2$.

Allo stesso modo di come abbiamo fatto con una compagnia assicurativa, possiamo calcolare la convenienza di una **lotteria** (intesa proprio come "il terno al lotto").

Al solito, schematizziamo: E è l'evento "vincita" (con probabilità p), d_1 la decisione "non gioco", d_2 la decisione "gioco". Il premio è la quantità S, e la quota per giocare è kS (in cui k è dell'ordine di p, ma un po' più grande per far sì che il banco vinca sempre qualcosa: $k > p$).

	d_1	d_2	$\mathbb{P}$
E	0	$S - kS$	p
$\overline{E}$	0	$-kS$	$1 - p$

Le speranze matematiche si calcolano usando la funzione utilità

$$\mathbb{E}(u(X_1)) = 0 \qquad \mathbb{E}(u(X_2)) < 0 \ \ (\text{perché } k > p)$$

per cui è abbastanza chiaro il messaggio: giocare non conviene mai.

Il fatto è che queste considerazioni tengono conto solo dei valori monetari, per cui era prevedibile che i risultati fossero questi. Chiedere se "conviene" giocare al lotto è più o meno come chiedere se "conviene" andare al cinema: perderemo sempre dei soldi, ma magari il film è bello.

8.3 Decisioni condizionate da una informazione: valore di una informazione

Schematizziamo una situazione complessa con molti eventi possibili $(E_1, \dots, E_m)$, ognuno con la sua probabilità, e molte decisioni possibili da prendere $(d_1, \dots, d_n)$, tramite una tabella:

	d_1	d_2	$\dots$	d_j	$\dots$	d_n	$\mathbb{P}$
E_1	x_{11}	x_{12}	$\dots$	x_{1j}	$\dots$	x_{1n}	p_1
E_2	x_{21}	x_{22}	$\dots$	x_{2j}	$\dots$	x_{2n}	p_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
E_i	x_{i1}	x_{i2}	$\dots$	x_{ij}	$\dots$	x_{in}	p_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
E_m	x_{m1}	x_{m2}	$\dots$	x_{mj}	$\dots$	x_{mn}	p_m

Guadagno atteso in caso di decisione d_j

$$\mathbb{E}(X_j) = \sum_{i=1}^{m} x_{ij} p_i$$

Tra tutte le decisioni possibili ci sarà quella che massimizza il guadagno; questa la chiamiamo "*decisione ottimale*". Il guadagno atteso con decisione ottimale lo indichiamo con

$$\overline{x} = \max_j \mathbb{E}(X_j)$$

Consideriamo ora la possibilità di ricevere una **perfetta informazione**, ovvero un'informazione che ci fa sapere con certezza quale degli eventi E_i si verificherà[1].

Il guadagno atteso in presenza di una perfetta informazione è

$$\overline{\overline{x}} = \sum_{i=1}^{m} p_i \ \max_j x_{ij}$$

Considerando che le speranze dei singoli eventi sono

$$\mathbb{E}(X_{j_0}) = \sum_{i=1}^{m} p_i \, x_{ij_0} \leq \sum_{i=1}^{m} p_i \ \max_j x_{ij} = \overline{\overline{x}}$$

il "valore" dell'informazione perfetta può essere calcolato come $V = \overline{\overline{x}} - \overline{x}$.

13/05/2014

Analizziamo ora il caso dell'informazione **parziale**: un'informazione $\mathcal{H}$, con risultati possibili $H_1, \ldots, H_r$, aventi probabilità $\rho_1, \ldots, \rho_r$.

Sono inoltre note $p_i^{(k)} = \mathbb{P}(E_i|H_k)$, che rappresentano i cambiamenti delle probabilità degli eventi E_i dati dalla conoscenza dell'informazione $\mathcal{H}$.

Il guadagno atteso nel caso in cui la risposta dell'informazione $\mathcal{H}$ sia H_k è calcolabile:

$$\mathbb{E}(X_j|H_k) = \sum_{i=1}^{m} x_{ij} p_i^{(k)}$$

$$\overline{x}^{(k)} = \max_j \ \mathbb{E}(X_j|H_k)$$

Da cui possiamo ricavare il guadagno atteso globalmente dall'informazione $\mathcal{H}$:

$$\overline{x}^{(\mathcal{H})} = \sum_{k=1}^{r} \rho_k \, \overline{x}^{(k)} = \sum_{k=1}^{r} \rho_k \ \max_j \ \mathbb{E}(X_j|H_k) = \sum_{k=1}^{r} \rho_k \ \max_j \left(\sum_{i=1}^{m} x_{ij} p_i^{(k)} \right)$$

[1] Attenzione: supponiamo di poter sapere esattamente quale evento accadrà, ma di non poter fare <u>nulla</u> per cambiarlo. Insomma, non stiamo facendo un film.

Proposizione 8.3.1 *Qualunque siano gli eventi e le probabilità, vale* $\overline{x} \leq \overline{x}^{(\mathcal{H})} \leq \overline{\overline{x}}$.[2]

Dimostrazione. Preso $1 \leq j_0 \leq n$ risulta

$$\overline{x}^{(\mathcal{H})} \geq \sum_k \rho_k \sum_i x_{ij_0}\, p_i^{(k)} = \sum_i x_{ij_0} \sum_k \rho_k\, p_i^{(k)}$$

ma $\rho_k = \mathbb{P}(H_k)$ e $p_i^{(k)} = \mathbb{P}(E_i|H_k)$, quindi $\sum_k \rho_k\, p_i^{(k)} = p_i$, da cui

$$\overline{x}^{(\mathcal{H})} \geq \sum_i x_{ij_0}\, p_i = \mathbb{E}(X_{j_0})$$

Visto che la disuguaglianza appena scritta vale per qualunque j_0 scelto, e che $\overline{x} = \max_j \mathbb{E}(X_j)$, allora $\overline{x} \leq \overline{x}^{(\mathcal{H})}$.

D'altra parte possiamo fare una maggiorazione diversa ed ottenere

$$\overline{x}^{(\mathcal{H})} \leq \sum_k \rho_k \sum_i \overline{x}_i\, p_i^{(k)} = \sum_i \overline{x}_i \sum_k \rho_k\, p_i^{(k)} = \overline{\overline{x}}$$

□

Esempio 8.3.1 Un tizio vuole comprare un terreno edificabile per 2 milioni di euro.

Poniamo che ci sia la possibilità che nel terreno ci siano reperti archeologici (evento E, con probabilità $p_1 = 0.45$), per cui lui non potrebbe più costruire, perdendo così 1 milione di euro. Il tizio ha due possibilità: non comprare (d_1) e investire i soldi in borsa, oppure comprare (d_2).

	d_1	d_2	$\mathbb{P}$
E	0.1	-1	0.45
$\overline{E}$	0.1	1	0.55
$\mathbb{E}$	0.1	0.1	

A parità di valore medio $\mathbb{E}$, se introduciamo una funzione utilità sicuramente il valore atteso ci dirà di non comprare, visto che la decisione di comprare ha la stessa media, ma varianza molto maggiore.

Supponiamo di fare un test (scavo preliminare) per scoprire quanto sia probabile che il tizio troverà dei reperti (risultato H) piuttosto che non li trovi (risultato $\overline{H}$).

Il test, esattamente come avevamo analizzato per le malattie, ha delle caratteristiche di specificità:

$$\mathbb{P}(H|E) = 0.8 \qquad \mathbb{P}(\overline{H}|\overline{E}) = 0.7$$

[2]Questa proposizione non ci deve meravigliare: quello che ci dice è che una informazione "perfetta" è meglio di una "parziale", che a sua volta è meglio di niente.

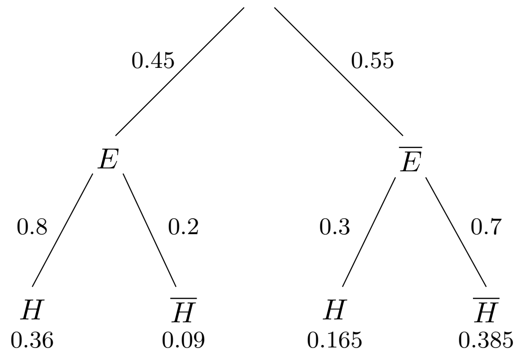
$$\mathbb{P}(\overline{H}|E) = 0.2 \qquad\qquad \mathbb{P}(H|\overline{E}) = 0.3$$

e delle probabilità sugli esiti $\rho_1 = \mathbb{P}(H)$, $\rho_2 = \mathbb{P}(\overline{H}) = 1 - \rho_1$.

Dopo aver fatto il test la probabilità dell'evento E (e del suo opposto) cambiano in

$$p_1^{(1)} = \mathbb{P}(E|H) \qquad\qquad p_2^{(1)} = \mathbb{P}(\overline{E}|H)$$

$$p_1^{(2)} = \mathbb{P}(E|\overline{H}) \qquad\qquad p_2^{(2)} = \mathbb{P}(\overline{E}|\overline{H})$$



Dal grafo calcoliamo le probabilità:

$$\mathbb{P}(H) = 0.525 = \rho_1 \qquad\qquad \mathbb{P}(\overline{H}) = 0.475 = \rho_2$$

$$p_1^{(1)} = \mathbb{P}(E|H) = \frac{\mathbb{P}(EH)}{\mathbb{P}(H)} = \frac{0.36}{0.525} = 0.686 \qquad\qquad p_2^{(1)} = \mathbb{P}(\overline{E}|H) = 1 - 0.686 = 0.314$$

$$p_1^{(2)} = \mathbb{P}(E|\overline{H}) = \frac{\mathbb{P}(E\overline{H})}{\mathbb{P}(\overline{H})} = \frac{0.09}{0.475} = 0.189 \qquad\qquad p_2^{(2)} = \mathbb{P}(\overline{E}|\overline{H}) = 1 - 0.189 = 0.811$$

Grazie a questi nuovi valori, la tabella iniziale diventa

	d_1	d_2
E	0.1	-1
$\overline{E}$	0.1	1
$\mathbb{E}(\cdot\|H)$	0.1	-0.372
$\mathbb{E}(\cdot\|\overline{H})$	0.1	0.622

Conclusione: il guadagno atteso assumendo $\mathcal{H}$ risulta essere $0.525 \cdot 0.1 + 0.475 \cdot 0.622$, che è sicuramente maggiore del guadagno atteso senza informazione.

Introduciamo all'interno del problema una funzione utilità

$$u(x) = 1 - e^{-\frac{x+1}{0.5}} = 1 - e^{-2(x+1)}$$

con cui possiamo calcolare i nuovi valori della tabella:

$$y_{11} = u(x_{11}) = u(0.1) = 0.8892 \qquad y_{12} = u(x_{12}) = u(-1) = 0$$

$$y_{21} = u(x_{21}) = u(0.1) = 0.8892 \qquad y_{22} = u(x_{22}) = u(1) = 0.9817$$

da cui ricaviamo la tabella, supponendo che il test abbia dato esito negativo $\overline{H}$:

	d_1	d_2	$\mathbb{P}$
E	0.8892	0	0.189
$\overline{E}$	0.8892	0.9817	0.811
$\mathbb{E}(\cdot\vert\overline{H})$	0.8892	0.7962	

L'utilità attesa con **perfetta informazione** è ben diversa:
La situazione di partenza è questa:

	d_1	d_2	$\mathbb{P}$
E	0.8892	0	0.45
$\overline{E}$	0.8892	0.9817	0.55

Se possiamo acquisire un'informazione certa ci comporteremo di conseguenza: con probabilità 0.45 decideremo di non comprare, avendo così un guadagno di 0.8892, mentre con probabilità 0.55 decideremo di comprare, avendo così un guadagno di 0.9817.

La media è dunque $0.45 \cdot 0.8892 + 0.55 \cdot 0.9817 = 0.9401$.

Per sapere il valore di una perfetta informazione dobbiamo calcolare alcuni valori:

$$x' = u^{-1}(0.8892) = 0.1$$

$$x'' = u^{-1}(0.9401) = 0.4075$$

$$\Longrightarrow \qquad x'' - x' = 0.3075$$

8.4 Programmazione Lineare: l'Algoritmo del Simplesso

Poniamo di essere dei ladri. Entriamo in un negozio per rubare dei televisori (TV), dei computer (PC), e dei telefoni (TEL), con le seguenti caratteristiche:
a_{11} = peso di una TV
a_{21} = volume di una TV
a_{31} = tempo per caricare una TV
a_{12} = peso di un PC
a_{22} = volume di un PC

a_{32} = tempo per caricare un PC
a_{13} = peso di un TEL
a_{23} = volume di un TEL
a_{33} = tempo per caricare un TEL
x_1 = numero di TV da prendere
x_2 = numero di PC da prendere
x_3 = numero di TEL da prendere

Sappiamo che dopo la rapina andremo dal nostro ricettatore per rivendere la merce rubata e guadagnare

$$p(x) = c_1x_1 + c_2x_2 + c_3x_3$$

Il nostro interesse è quindi massimizzare $p(x)$ con dei vincoli sulle quantità di partenza, date dalla grandezza del furgone usato per la rapina e del tempo a disposizione prima dell'arrivo della polizia:

peso: $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \leq b_1$
volume: $a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \leq b_2$
tempo: $a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \leq b_3$

16/05/2014

Più in generale, lasciando da parte le rapine e quello che ne consegue, la programmazione lineare si occupa di risolvere problemi di questo tipo: trovare $(x_1, \ldots, x_n)$, con $x_j \geq 0 \ \forall j$ tali che verificano

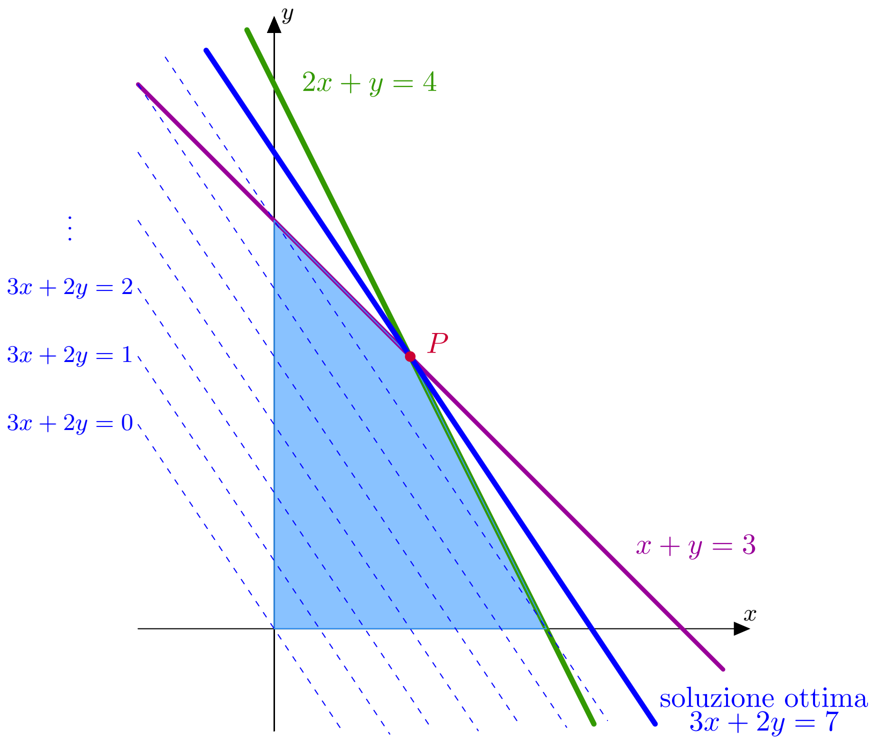
$$\circledast \quad \sum_{j=1}^{n} a_{ij}x_j = b_i \qquad \text{oppure} \qquad \sum_{j=1}^{n} a_{ij}x_j \leq b_i \qquad i = 1, \ldots, m$$

e sia massimo il valore di

$$p(x) = \sum_{j=1}^{n} c_jx_j$$

Esempio 8.4.1 Vogliamo massimizzare la funzione $p(x, y) = 3x+2y$ con i seguenti vincoli:

$$\begin{cases} x + y \leq 3 \\ 2x + y \leq 4 \end{cases}$$



La soluzione ottimale del problema sta nel punto di intersezione dei due vincoli: $(x, y) = (1, 2)$. Il valore massimo di p possiamo dunque calcolarlo facilmente: $p(x, y) = p(1, 2) = 3 + 4 = 7$.

Osservazione Ogni vincolo del tipo $\sum a_{ij}x_j \leq b_i$ possiamo riscriverlo nella forma

$$\sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i \qquad \text{con } x_{n+1} \geq 0$$

introducendo le "*variabili di scarto*" x_{n+i}, così da riportare ogni vincolo di tipo "$\leq$" in uno di tipo "$=$".

8.4.1 Ipotesi fondamentali

Introduciamo i termini dei problemi di programmazione lineare:
Programma=soluzione del problema $\circledast$.
Programma ammissibile = soluzione di $\circledast$ con ogni $x_j \geq 0$.
Funzione oggetto (o obiettivo) = $p(x)$.
Regione ammissibile = $\{$ soluzioni di $\circledast$ con $x_j \geq 0\}$.
Programma ottimale = programma ammissibile che rende massima $p(x)$ nella regione ammissibile.

Troviamo un algoritmo che ci permetta di arrivare alla soluzione ottimale che cerchiamo. Trasformiamo innanzitutto il sistema di equazioni[3] (i vincoli) in forma matriciale:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \qquad A_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

$$\Longrightarrow \qquad \sum_{j=1}^{n} x_j A_j = B$$

Vi sono alcune condizioni da porre affinché il problema sia ben posto:

1. $A^* = Span(A_1, \dots, A_n)$ ha dimensione[4] $dim A^* = k$.
2. $B \in A^*$.
3. $\{A_{v_1}, \dots, A_{v_k}\}$ base di A^*.
4. Per le colonne di A: per ogni $1 \leq j \leq n$ vale
$$A_j = \sum_{r=1}^{k} \alpha_{rj} A_{v_r}$$
5. Per il vettore B vale
$$B = \sum_{r=1}^{k} \beta_r A_{v_r}$$
6. $\beta_r \geq 0$, $r = 1, 2, \dots, k$.

[3]Ricordando che nel caso alcuni vincoli siano espressi in termini di disequazioni possiamo sempre introdurre le variabili di scarto e riportare tutto a delle equazioni.
[4]Visto come sottospazio vettoriale di $\mathbb{R}^m$

Osservazione L'ipotesi numero 6 implica che la regione ammissibile non è vuota.

Possiamo infatti trovare subito una soluzione ammissibile:

$$\overline{x}_j = \begin{cases} \beta_r & \text{se } j = v_r,\ r = 1, \dots, k \\ 0 & \text{altrimenti} \end{cases}$$

che soluzione è? Supponiamo di avere 5 colonne A_1, A_2, A_3, A_4, A_5, e che la base per lo spazio A^* sia data da $A_3 = A_{v_1}$ e $A_5 = A_{v_2}$. Allora il vettore $\overline{x}$ definito sopra è così scritto:

$$\overline{x} = \begin{pmatrix} 0 \\ 0 \\ \beta_1 \\ 0 \\ \beta_2 \end{pmatrix}$$

ed è effettivamente una soluzione ammissibile:

$$\sum_j \overline{x}_j A_j = \sum_{r=1}^{k} \overline{x}_{v_r} A_{v_r} = \sum_{r=1}^{k} \beta_r A_{v_r} = B$$

8.4.2 La tabella del simplesso

Introduciamo ora l'elemento più importante, su cui si basa totalmente l'algoritmo che stiamo spiegando: la **tabella del Simplesso**.

		A_1	A_2	$\dots$	A_s	$\dots$	A_n	B
x_{v_1}	c_{v_1}	α_{11}	α_{12}	$\dots$	α_{1s}	$\dots$	α_{1n}	β_1
x_{v_2}	c_{v_2}	α_{21}	α_{22}	$\dots$	α_{2s}	$\dots$	α_{2n}	β_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{v_r}	c_{v_r}	α_{r1}	α_{r2}	$\dots$	α_{rs}	$\dots$	α_{rn}	β_r
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{v_k}	c_{v_k}	α_{k1}	α_{k2}	$\dots$	α_{ks}	$\dots$	α_{kn}	β_k
		$z_1 - c_1$	$z_2 - c_2$	$\dots$	$z_s - c_s$	$\dots$	$z_n - c_n$	z

in cui z_j e $z = p(\overline{x})$ sono ottenuti con i seguenti calcoli:

$$z_j = \sum_{r=1}^{k} \alpha_{rj} c_{v_r} \qquad z = \sum_{r=1}^{k} \beta_r c_{v_r}$$

e i c_{v_i} sono i coefficienti delle variabili x_{v_i} della funzione $p(x)$ da massimizzare.

Si capisce tutto meglio con degli esempi. Vediamone uno per iniziare, poi ne vedremo molti altri.

Esempio 8.4.2 Poniamo che $A^* = Span\{A_1, A_2, A_3, A_4, A_5\}$, con $k = 2$; scegliamo come base=$\{A_{v_1}, A_{v_2}\} = \{A_3, A_5\}$, e supponiamo di dover massimizzare la funzione $p(x) = 3x_1 - 7x_2 + x_3 - x_4 + 8x_5$, all'interno dei vincoli:

$$\begin{cases} 3x_1 - 5x_2 + 8x_3 + x_4 - x_5 = 3 \\ 7x_1 + 8x_2 + 9x_3 + 10x_4 + 11x_5 = 12 \end{cases}$$

I valori da scrivere nella tabella sono $x_{v_1} = x_3 \Rightarrow c_{v_1} = 1$, $x_{v_2} = x_5 \Rightarrow c_{v_2} = 8$.
Portiamo i vincoli in forma matriciale:

$$A = \begin{pmatrix} 3 & -5 & 8 & 1 & -5 \\ 7 & 8 & 9 & 10 & 11 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 3 \\ 7 \end{pmatrix} \qquad \dots \qquad A_5 = \begin{pmatrix} -5 \\ 11 \end{pmatrix} \qquad B = \begin{pmatrix} 3 \\ 12 \end{pmatrix}$$

Costruiamo quindi la tabella:

		A_1	A_2	A_3	A_4	A_5	B
x_3	1	α_{11}	α_{12}	1	α_{14}	0	β_1
x_5	8	α_{21}	α_{22}	0	α_{24}	1	β_2
		$z_1 - 3$	$z_2 + 7$	$1 - 1$	$z_4 + 1$	$8 - 8$	$\beta_1 + 8\beta_2$

Sperando che sia un po' più chiaro, torniamo alla teoria: sappiamo già che $\overline{x}$ (definita sopra) è una soluzione ammissibile e ha un guadagno pari a $z = p(\overline{x})$. La domanda è se si può trovare un valore maggiore per $p(x)$ oppure no.

Ci sono tre situazioni distinte da gestire:

a) $z_j - c_j \geq 0$, $\forall j \in \{1, \dots, n\}$.

$$\overline{x} = \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} \qquad \overline{x}_j = \begin{cases} \beta_r & \text{se } j = v_r \\ 0 & \text{altrimenti} \end{cases}$$

per cui $\overline{x}$ è un programma ottimale e abbiamo trovato il valore ottimale $z = p(\overline{x})$.

b) $\exists\, s$ per cui $z_s - c_s < 0$; inoltre $\alpha_{rs} \leq 0$ per $r = 1, \dots, k$.

$p(x)$ è superiormente illimitata nella regione ammissibile.

c) $\exists\, s$ per cui $z_s - c_s < 0$; inoltre $\forall j$ tale che $z_j - c_j < 0$, $\exists\, r$ tale che $\alpha_{rj} > 0$.

Si cambia la base di A^*, e si ricompila la tabella relativamente alla nuova base.

Problema del metodo: c'è la possibilità che il metodo entri in loop, ma i dati in entrata devono essere scelti in maniera molto molto accorta, per cui in applicazioni pratiche la probabilità che ciò accada è trascurabile.

Teorema 8.4.1 *Supponiamo le ipotesi 1−5 di sopra. Allora vale*

$$(1) \quad \sum_{j=1}^{n} x_j A_j = B \quad \Longleftrightarrow \quad \sum_{j=1}^{n} \alpha_{rj} x_j = \beta_r \quad 1 \le r \le k \quad (2)$$

Dimostrazione.

$$(1) \quad \Longleftrightarrow \quad \sum_{j=1}^{n} \left(x_j \cdot \sum_{r=1}^{k} \alpha_{rj} A_{v_r} \right) = \sum_{r=1}^{k} \beta_r A_{v_r} \quad \Longleftrightarrow$$

$$\Longleftrightarrow \quad \sum_{r=1}^{k} \left(\sum_{j=1}^{n} \alpha_{rj} x_j \right) A_{v_r} = \sum_{r=1}^{k} \beta_r A_{v_r} \quad \Longleftrightarrow \quad (2)$$

□

18/05/2014

Nota: in tutti i prossimi teoremi si supporranno le ipotesi $1-5$ di sopra, quindi non lo ripeteremo.

Teorema 8.4.2 *Per ogni* $x = (x_1, \ldots, x_n)$ *programma ammissibile risulta*

$$p(x) = z - \sum_{j \in \mathcal{J}} (z_j - c_j) x_j$$

in cui $\mathcal{J} = \{1, 2, \ldots, n\} \backslash \{v_1, \ldots, v_k\}$.

Dimostrazione. Dalle ipotesi (4) e (5), ricaviamo

$$A_{v_r} = \sum_{l=1}^{k} \alpha_{l,v_r} A_{v_l} \qquad\qquad \alpha_{l,v_r} = \begin{cases} 0 & \text{se } l \ne r \\ 1 & \text{se } l = r \end{cases}$$

Dalla relazione (2) del teorema precedente ricaviamo

$$x_{v_r} = \beta_r - \sum_{j \in \mathcal{J}} \alpha_{rj} x_j \qquad 1 \le r \le k \quad (\circledast)$$

Sia quindi x soddisfacente $\circledast$; Per questo x vale

$$p(x) = \sum_{j=1}^{n} c_j x_j = \sum_{r=1}^{k} c_{v_r} x_{v_r} + \sum_{j \in \mathcal{J}} c_j x_j =$$

$$= \sum_{r=1}^{k} c_{v_r} \left(\beta_r - \sum_{j \in \mathcal{J}} \alpha_{rj} x_j \right) + \sum_{j \in \mathcal{J}} c_j x_j =$$

$$= z - \sum_{r=1}^{k} c_{v_r} \sum_{j \in \mathcal{J}} \alpha_{rj} x_j + \sum_{j \in \mathcal{J}} c_j x_j = z - \sum_{j \in \mathcal{J}} \left(\sum_{r=1}^{k} c_{v_r} \alpha_{rj} - c_j \right) x_j =$$

$$= z - \sum_{j \in \mathcal{J}} (z_j - c_j) x_j$$

□

8.4.3 Analisi delle situazioni

Adesso iniziamo ad analizzare, tramite dei teoremi, le tre situazioni possibili dell'algoritmo del simplesso, che abbiamo descritto qualche pagina fa. Prima ci eravamo limitati a dire cosa avremmo dovuto fare, adesso dimostriamolo sul serio.

Teorema 8.4.3 (Situazione a) *Nella situazione*[5] (a) *sopra descritta, ovvero se vale* $z_j - c_j > 0$, *per* $j = 1, 2, \ldots, n$, *allora* $\overline{x}$, *definito come*

$$\overline{x}_j = \begin{cases} \beta_r & \textit{se } j = v_r \\ 0 & \textit{altrimenti} \end{cases}$$

è un programma ottimale.

Dimostrazione. Sappiamo che $p(\overline{x}) = z$, quindi quello che dobbiamo dimostrare è che, preso $x = (x_1, \ldots, x_n)$ un altro programma ammissibile, valga $p(x) < p(\overline{x}) = z$.

Questo si fa abbastanza semplicemente:

$$p(x) = z - \sum_{j \in \mathcal{J}} \underbrace{(z_j - c_j)}_{\geq 0} \underbrace{x_j}_{\geq 0} \leq z$$

□

[5] vedi lezione del 16/05/2014

Teorema 8.4.4 (Situazione b) *Se siamo nella situazione*[6] (b)*, allora* $p(x)$ *è superiormente illimitata nella regione ammissibile.*

Dimostrazione. La situazione (b) è descritta dalle relazioni: $\exists\, s$ per cui $z_s - c_s < 0$; inoltre $\alpha_{rs} \leq 0$ per $r = 1, \dots, k$.

Sia dunque $x(t)$, $t \in \mathbb{R}$ il vettore con le seguenti componenti:

$$x_j = \begin{cases} t & \text{se } j = s \\ \beta_r - t\alpha_{rs} & \text{se } j = v_r, \quad r = 1, \dots, k \\ 0 & \text{altrimenti} \end{cases}$$

$x(t)$ è un programma ammissibile $\forall t \geq 0$, infatti le sue componenti non nulle sono sicuramente positive, inoltre

$$\sum_{j=1}^{n} x_j A_j = tA_s + \sum_{r=1}^{k} (\beta_r - t\alpha_{rs}) A_{v_r} =$$

$$= \underbrace{\sum_{r=1}^{k} \beta_r A_{v_r}}_{B} + tA_s - t \underbrace{\sum_{r=1}^{k} \alpha_{rs} A_s}_{A_s} = B$$

Quindi $p(x(t)) = z - t(z_s - c_s) \xrightarrow{t \to +\infty} +\infty$, poiché $z_s - c_s < 0$, quindi la regione ammissibile è illimitata perché esistono infinite soluzioni $x(t)$ ammissibili.

□

Situazione c:[7] in quest'ultima situazione dobbiamo cambiare la base di vettori A_{v_i} per poter scegliere una soluzione migliore di quella trovata. Vediamo quale cambio di base fare e perché ci porta ad una soluzione migliore:

Proposizione 8.4.1 *Sostituendo* A_{v_r} *con* A_s*, si ottiene una nuova base di* $A^* \iff \alpha_{rs} \neq 0$.

Dimostrazione. Trasformiamo la base: $\{A_{v_1}, \dots, A_{v_{r-1}}, \overset{A_s}{\cancel{A_{v_r}}}, A_{v_{r+1}}, \dots, A_{v_k}\}$, e vediamo quando questa è davvero una nuova base, ovvero indaghiamo l'indipendenza lineare dei vettori.

$$\xi_r A_s + \sum_{\substack{l \neq r \\ l=1}}^{k} \xi_l A_{v_l} = \xi_r \sum_{\substack{l \neq r \\ l=1}}^{k} \alpha_{ls} A_{v_l} + \xi_r \alpha_{rs} A_{v_r} + \sum_{\substack{l \neq r \\ l=1}}^{k} \xi_l A_{v_l} =$$

$$= \xi_r \alpha_{rs} A_{v_r} + \sum_{l \neq r} (\xi_r \alpha_{ls} + \xi_l) A_{v_l} = \underline{0} \iff \begin{cases} \xi_r \alpha_{rs} = 0 \\ \xi_r \alpha_{ls} + \xi_l = 0 \quad \forall l \neq r \end{cases}$$

[6] vedi lezione del 16/05/2014
[7] vedi lezione del 16/05/2014

Se $\alpha_{rs} \neq 0$, affinché i vettori siano indipendenti dovremo prendere $\xi_r = 0$ e anche $\xi_l = 0$ per $l \neq r$.

Se, invece, $\alpha_{rs} = 0$, affinché i vettori siano indipendenti dovremo prendere $\xi_r = 1$, e troviamo $\xi_l = -\alpha_{ls}$ per $l \neq r$.

$\square$

Il cambiamento di base che abbiamo descritto nella proposizione è chiamato **Trasformazione Pivotale**. Vediamo meglio come funziona: nella vecchia base avevamo

$$A_j = \sum_{l=1}^{k} \alpha_{lj} A_{v_l}$$

mentre nella nuova base (in cui abbiamo sostituito il vettore r-esimo A_{v_r} con A_s)

$$A_j = \sum_{\substack{l \neq r \\ l=1}}^{k} \alpha'_{lj} A_{v_l} + \alpha'_{rj} A_s \quad (\circledast)$$

per capire cosa è cambiato scriviamo i nuovi coefficienti α' in funzione dei vecchi α:

$$\begin{aligned} \circledast \quad A_j &= \sum_{l \neq r} \alpha'_{lj} A_{v_l} + \alpha'_{rj} \sum_{\substack{l \neq r \\ l=1}}^{k} \alpha_{ls} A_{v_l} + \alpha'_{rj} \alpha_{rs} A_{v_r} = \\ &= \alpha'_{rj} \alpha_{rs} A_{v_r} + \sum_{l \neq r} \left(\alpha'_{lj} + \alpha'_{rj} \alpha_{ls} \right) A_{v_l} \end{aligned}$$

Abbiamo due scritture del vettore A_j; visto che devono essere uguali otteniamo

$$\begin{cases} \alpha'_{rj} \alpha_{rs} = \alpha_{rj} \\ \alpha'_{lj} + \alpha'_{rj} \alpha_{ls} = \alpha_{lj} \quad (\forall l \neq r) \end{cases} \Longrightarrow \begin{cases} \alpha'_{rj} = \frac{\alpha_{rj}}{\alpha_{rs}} \\ \alpha'_{lj} = \alpha_{lj} - \frac{\alpha_{rj}}{\alpha_{rs}} \alpha_{ls} \quad (\forall l \neq r) \end{cases}$$

Osservazione Dalle relazioni appena trovate possiamo ricavare $\alpha'_{rs} = 1$. Non solo, per $l \neq r$ vale $\alpha'_{ls} = \alpha_{ls} - 1 \cdot \alpha_{ls} = 0$.

Per B possiamo ripetere lo stesso ragionamento:

$$B = \sum_{l=1}^{k} \beta_l A_{v_l} \qquad\qquad B = \sum_{l \neq r} \beta'_l A_{v_l} + \beta'_r A_s$$

$$\Longrightarrow \quad B = \sum_{l \neq r} \beta'_l A_{v_l} + \beta'_r \left(\sum_{l \neq r} \alpha_{ls} A_{v_l} + \alpha_{rs} A_{v_r} \right) = \beta'_r \alpha_{rs} A_{v_r} + \sum_{l \neq r} \left(\beta'_l + \beta'_r \alpha_{ls} \right) A_{v_l}$$

da cui otteniamo le componenti

$$\begin{cases} \beta'_r \alpha_{rs} = \beta_r \\ \beta'_l + \beta'_r \alpha_{ls} = \beta_l \quad (\forall l \neq r) \end{cases} \implies \begin{cases} \beta'_r = \frac{\beta_r}{\alpha_{rs}} \\ \beta'_l = \beta_l - \frac{\beta_r}{\alpha_{rs}} \alpha_{ls} \quad (\forall l \neq r) \end{cases}$$

20/05/2014

Adesso che abbiamo chiarito la "trasformazione pivotale", vediamo operativamente cosa fare nella situazione *c*:[8] "esce" il vettore A_{v_r}, "entra" il vettore A_s (cambio di base sopra descritto). Nella nuova tabella del simplesso abbiamo i coefficienti cambiati come avevamo scoperto poco fa:

$$\begin{cases} \alpha'_{rj} = \frac{\alpha_{rj}}{\alpha_{rs}} \\ \alpha'_{lj} = \alpha_{lj} - \frac{\alpha_{rj}}{\alpha_{rs}} \alpha_{ls} \quad (\forall l \neq r) \end{cases} \qquad \begin{cases} \beta'_r = \frac{\beta_r}{\alpha_{rs}} \\ \beta'_l = \beta_l - \frac{\beta_r}{\alpha_{rs}} \alpha_{ls} \quad (\forall l \neq r) \end{cases}$$

La vecchia tabella del simplesso M si trasforma in $M' = H \cdot M$, in cui H è la matrice seguente:

$$H = \left(\begin{array}{c|cccccccc} & 1 & 2 & 3 & & r & & k \\ \hline 1 & 1 & 0 & 0 & \dots & -\frac{\alpha_{1s}}{\alpha_{rs}} & \dots & 0 \\ 2 & 0 & 1 & 0 & \dots & -\frac{\alpha_{2s}}{\alpha_{rs}} & \dots & 0 \\ & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ r & 0 & 0 & 0 & \dots & \frac{1}{\alpha_{rs}} & \dots & 0 \\ & \vdots & \vdots & \vdots & & \vdots & & \vdots \\ k & 0 & 0 & 0 & \dots & -\frac{\alpha_{ks}}{\alpha_{rs}} & \dots & 1 \end{array}\right)$$

Teorema 8.4.5 *Sia $\alpha_{rs} \neq 0$, e sia*

$$\frac{\beta_r}{\alpha_{rs}} = \min \left\{ \frac{\beta_l}{\alpha_{ls}} \;\middle|\; l \in \{1, \dots, k\},\ \alpha_{ls} > 0 \right\}$$

Allora $\beta'_l \geq 0$ per $l = 1, 2, \dots, k$ (ovvero la nuova base di A è associata ad un programma ammissibile).

Dimostrazione. Visto che $\alpha_{rs} > 0 \implies \beta'_r \geq 0$.

Se $l \neq r$, $\beta'_l = \beta_l - \frac{\alpha_{ls}}{\alpha_{rs}} \beta_r$, quindi i casi possibili sono due:

1. Se $\alpha_{ls} \leq 0$, allora $\beta'_l \geq 0$

2. Se $\alpha_{ls} > 0$, allora

$$\beta'_l = \alpha_{ls} \left(\frac{\beta_l}{\alpha_{ls}} - \frac{\beta_r}{\alpha_{rs}} \right) \geq 0$$

[8] vedi lezione del 16/05/2014

□

Teorema 8.4.6 *Se $\alpha_{rs} \neq 0$ e $z_s - c_s < 0$, allora $z' \geq z$.*[9]

Inoltre $z' = z \iff \beta_r = 0$, e in tale caso $\overline{x}' = \overline{x}$.

Dimostrazione.

$$z' - z = \overbrace{c_s\beta_r' + \sum_{l\neq r} c_{v_l}\beta_l'}^{z'} - \overbrace{\sum_{l\neq r} c_{v_l}\beta_l - c_{v_r}\beta_r}^{z} =$$

$$= c_s\frac{\beta_r}{\alpha_{rs}} + \sum_{l\neq r} c_{v_l}\beta_l - \sum_{l\neq r} c_{v_l}\frac{\alpha_{ls}}{\alpha_{rs}}\beta_r - \sum_{l\neq r} c_{v_l}\beta_l - c_{v_r}\beta_r =$$

$$= c_s\frac{\beta_r}{\alpha_{rs}} - \sum_{l\neq r} c_{v_l}\frac{\alpha_{ls}}{\alpha_{rs}}\beta_r - c_{v_r}\beta_r = \frac{\beta_r}{\alpha_{rs}}\left(c_s - \sum_{l\neq r} c_{v_l}\alpha_{ls} - c_{v_r}\alpha_{rs}\right) =$$

$$= \frac{\beta_r}{\alpha_{rs}}\left(c_s - \sum_{l=1}^{k} c_{v_l}\alpha_{ls}\right) = \frac{\beta_r}{\alpha_{rs}}\left(c_s - z_s\right) \geq 0$$

□

8.4.4 Esempi pratici

Ora che abbiamo analizzato il funzionamento dell'algoritmo del simplesso, e la teoria che sta dietro al suo funzionamento, vediamo alcuni esempi pratici per capire davvero come applicarlo.

Nell'algoritmo del simplesso ci sono due criteri fondamentali che abbiamo analizzato e che dovremo usare:

Criterio di entrata: entra un A_s per cui $z_s - c_s < 0$

Criterio di uscita: esce A_{v_r}, con $\alpha_{rs} > 0$ e

$$\frac{\beta_r}{\alpha_{rs}} = \min\left\{\frac{\beta_l}{\alpha_{ls}} \mid l \in \{1, \ldots, k\},\ \alpha_{ls} > 0\right\}$$

Nota: questi criteri non sono univoci, potremmo trovare più vettori che li verificano. In questo caso la scelta è abbastanza libera.

[9] Qui z' è il vettore z della matrice del simplesso dopo la trasformazione detta sopra.

Esempio 8.4.3 Cerchiamo $\max p(x_1, x_2) = 5x_1 + 3x_2$ sotto i seguenti vincoli

$$\begin{cases} x_1 + x_2 \leq 400 \\ 2x_1 + x_2 \leq 625 \end{cases}$$

Nell'ottica di applicare l'algoritmo del simplesso introduciamo delle "variabili di scarto" che saturino (appunto) gli scarti, così da trasformare i vincoli in

$$\begin{cases} x_1 + x_2 + x_3 = 400 \\ 2x_1 + x_2 + x_4 = 625 \end{cases} \longrightarrow \begin{array}{cccc} \left(\begin{array}{cccc} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \\ \begin{array}{cccc} A_1 & A_2 & A_3 & A_4 \end{array} \end{array}$$

Dobbiamo ora scegliere due vettori A_i che formino una base per $\mathbb{R}^2$. Visti come sono fatti gli A_i potremmo scegliere una qualsiasi coppia di vettori, visto che sono tutte linearmente indipendenti, ma visto che c'è la possibilità di scegliere la base canonica di $\mathbb{R}^2$ (A_3, A_4) viene abbastanza naturale scegliere quella.

Prendiamo quindi $A_{v_1} = A_3$, $A_{v_2} = A_4$.

Scriviamo la matrice del simplesso:

				A_{v_1}	A_{v_2}	
		A_1	A_2	A_3	A_4	B
x_3	0	1	1	1	0	400
x_4	0	2	1	0	1	625
		-5	-3	0	0	0

Nota: i coefficienti vengono esattamente come quelli del sistema proprio perché abbiamo preso la base canonica! Altrimenti non sarebbe stato così.

Abbiamo quindi $\overline{x} = (0, 0, 400, 625)$. Questa soluzione non ha senso perché si "scarica" tutto sulle variabili di scarto.

Nell'esempio del ladro che sta facendo una rapina vorrebbe dire che il ladro non ha preso niente...

Quindi sicuramente questa non è la soluzione che massimizza il guadagno. Continuiamo con l'algoritmo.

Siamo nella situazione c ($\exists\, s$ tale che $z_s - c_s < 0$), e abbiamo due vettori "entranti" possibili: A_1 e A_2. Decidiamo di prendere A_1 perché ha valore assoluto maggiore.

Vettore entrante: A_1 $(s = 1)$.

Il vettore uscente lo calcoliamo, ricordando che dobbiamo prendere il vettore A_{v_r} tale che sia minimo il valore di $\frac{\beta_r}{\alpha_{rs}}$ (vedi teoremi):

$$\frac{\beta_1}{\alpha_{1s}} = \frac{\beta_1}{\alpha_{11}} = 400 \qquad \frac{\beta_2}{\alpha_{2s}} = \frac{\beta_2}{\alpha_{21}} = \frac{625}{2} < 400$$

Vettore uscente: $A_{v_2} = A_4$.

Calcoliamo la nuova matrice; vogliamo trasformare la matrice in questo modo:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 400 \\ 2 & 1 & 0 & 1 & 625 \\ \uparrow & & \circlearrowleft & \downarrow & \\ \textit{entra} & & \textit{resta} & \textit{esce} & \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & * & 1 & * & * \\ 1 & * & 0 & * & * \end{pmatrix}$$

La trasformazione deve avvenire tramite operazioni di riga e colonna, facciamo i passaggi:

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 400 \\ 2 & 1 & 0 & 1 & 625 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & 1 & 0 & 400 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{625}{2} \end{pmatrix} \longrightarrow$$

$$\longrightarrow \begin{pmatrix} 0 & \frac{1}{2} & 1 & -\frac{1}{2} & \frac{175}{2} \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{625}{2} \end{pmatrix}$$

Quindi la nuova tabella del simplesso diventa questa:

		A_{v_2}		A_{v_1}		
		A_1	A_2	A_3	A_4	B
x_3	0	0	$\frac{1}{2}$	1	$-\frac{1}{2}$	$\frac{175}{2}$
x_1	5	1	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{625}{2}$
		0	$-\frac{1}{2}$	0	$\frac{5}{2}$	$\frac{3125}{2}$

Siamo di nuovo in situazione c, visto che abbiamo $z_2 - c_2 < 0$.
Ripetiamo il ragionamento di prima:

Vettore entrante: A_2 $(s = 2)$.

Per il vettore uscente facciamo i calcoli

$$\frac{\beta_1}{\alpha_{12}} = \frac{\frac{175}{2}}{\frac{1}{2}} = 175 \qquad \frac{\beta_2}{\alpha_{22}} = \frac{\frac{625}{2}}{\frac{1}{2}} = 625$$

prendiamo il più piccolo, per cui il **vettore uscente** è $A_{v_1} = A_3$.

Calcoliamo la nuova matrice; vogliamo trasformare la matrice in questo modo:

$$\begin{array}{ccccc} \left(\begin{array}{c} 0 \\ 1 \end{array}\right. & \begin{array}{c} \frac{1}{2} \\ \frac{1}{2} \end{array} & \begin{array}{c} 1 \\ 0 \end{array} & \begin{array}{c} -\frac{1}{2} \\ \frac{1}{2} \end{array} & \left.\begin{array}{c} \frac{175}{2} \\ \frac{625}{2} \end{array}\right) \\ \circlearrowleft & \uparrow & \downarrow & & \\ \textit{resta} & \textit{entra} & \textit{esce} & & \end{array} \longrightarrow \begin{pmatrix} 0 & 1 & * & * & * \\ 1 & 0 & * & * & * \end{pmatrix}$$

Facciamo i calcoli (stavolta tutti in una volta):

$$\begin{pmatrix} 0 & \frac{1}{2} & 1 & -\frac{1}{2} & \frac{175}{2} \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{625}{2} \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 & 2 & -1 & 175 \\ 1 & 0 & -1 & 1 & 225 \end{pmatrix}$$

Ricompiliamo la tabella del simplesso:

		A_{v_2} A_1	A_{v_1} A_2	A_3	A_4	B
x_2	3	0	1	2	-1	175
x_1	5	1	0	-1	1	225
		0	0	1	2	1650

Siamo finalmente nella situazione a ($z_i - c_i \geq 0\ \forall i$), quindi abbiamo trovato una soluzione ottimale:

$$\overline{x} = (225, 175, 0, 0)$$

In questo caso le risorse vengono impiegate interamente, perché le variabili di scarto sono entrambe nulle, e il guadagno totale è dato da $p(\overline{x}) = z = 1650$.

Esempio 8.4.4 Data la funzione $p(x_1, x_2) = x_1 + 2x_2$, vogliamo massimizzarla nei vincoli

$$\begin{cases} 4x_1 + x_2 \leq 10 \\ 9x_1 + x_2 \leq 20 \end{cases}$$

Come al solito, trasformiamo i vincoli introducendo le variabili di scarto:

$$\begin{cases} 4x_1 + x_2 + x_3 = 10 \\ 9x_1 + x_2 + x_4 = 20 \end{cases}$$

e scriviamo dunque i vettori della base

$$\begin{cases} 4x_1 + x_2 + x_3 = 10 \\ 9x_1 + x_2 + x_4 = 20 \end{cases} \longrightarrow \begin{array}{c} \begin{pmatrix} 4 & 1 & 1 & 0 & 10 \\ 9 & 1 & 0 & 1 & 20 \end{pmatrix} \\ \begin{array}{ccccc} A_1 & A_2 & A_3 & A_4 & B \end{array} \end{array}$$

Scegliamo due vettori che facciano da base: $A_{v_1} = A_3$, $A_{v_2} = A_4$, per le stesse motivazioni dell'esempio precedente.

Scriviamo quindi la tabella del simplesso:

				A_{v_1}	A_{v_2}	
		A_1	A_2	A_3	A_4	B
x_3	0	4	1	1	0	10
x_4	0	9	1	0	1	20
		-1	-2	0	0	0

Abbiamo due indici con $z_s - c_s < 0$; Scegliamo di prediligere l'indice 2 perché ha valore assoluto maggiore.

Vettore entrante: $A_s = A_2$ $(s = 2)$.

Per il vettore uscente svolgiamo i calcoli:

$$\frac{\beta_1}{\alpha_{12}} = 10 \qquad\qquad \frac{\beta_2}{\alpha_{22}} = 20$$

prendiamo il minimo, quindi il **vettore uscente** è $A_{v_1} = A_3$.

$$\begin{array}{c} \begin{pmatrix} 4 & 1 & 1 & 0 & 10 \\ 9 & 1 & 0 & 1 & 20 \end{pmatrix} \\ \begin{array}{ccccc} & \uparrow & \downarrow & \circlearrowleft & \\ & \textit{entra} & \textit{esce} & \textit{resta} & \end{array} \end{array} \longrightarrow \begin{pmatrix} * & 1 & * & 0 & * \\ * & 0 & * & 1 & * \end{pmatrix}$$

Facendo i calcoli:

$$\begin{pmatrix} 4 & 1 & 1 & 0 & 10 \\ 9 & 1 & 0 & 1 & 20 \end{pmatrix} \longrightarrow \begin{pmatrix} 4 & 1 & 1 & 0 & 10 \\ 5 & 0 & -1 & 1 & 10 \end{pmatrix}$$

Quindi la nuova tabella del simplesso diventa questa:

			A_{v_1}		A_{v_2}	
		A_1	A_2	A_3	A_4	B
x_2	2	4	1	1	0	10
x_4	0	5	0	-1	1	10
		7	0	2	0	20

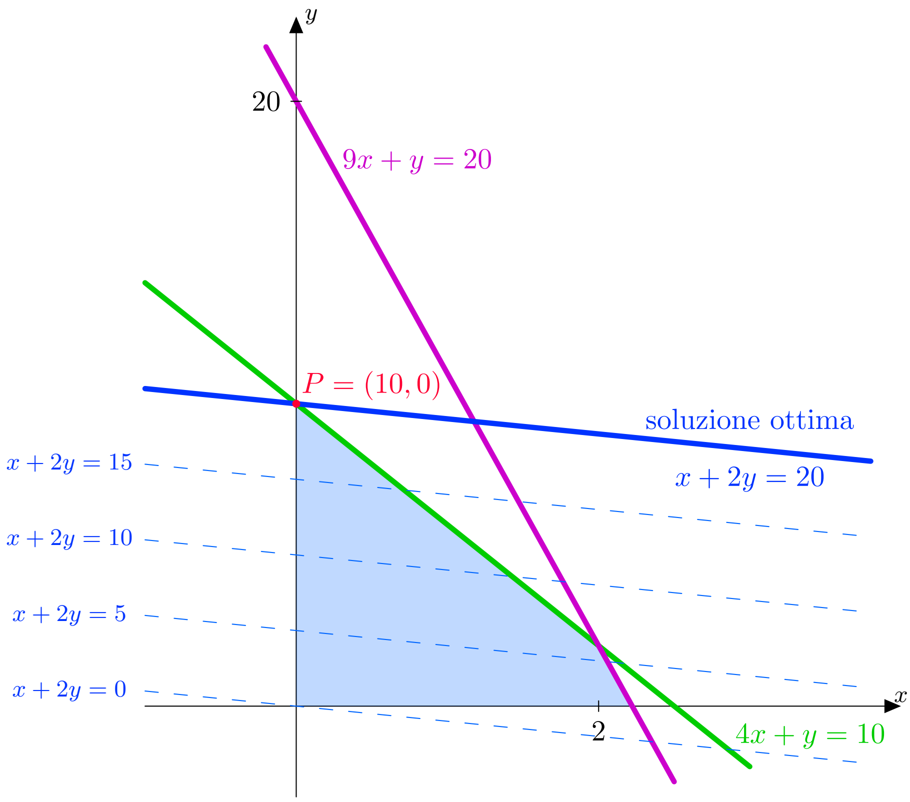
Abbiamo quindi già ottenuto una soluzione ottimale (poiché $z_i - c_i \geq 0\ \forall i$), che è

$$\overline{x} = (0, 10, 0, 10)$$

Nota: La soluzione trovata, seppur sicuramente ottimale, ci impone di usare solo una delle due risorse: usiamo 10 unità di x_2 e 0 di x_1, con una rimanenza di 10 nella variabile di scarto x_4.

Non è strano che sia davvero una soluzione ottimale?

In realtà no; la situazione si può spiegare anche tramite un grafico:



Come si può vedere dal grafico con gli assi non in scala, la soluzione che abbiamo trovato equivale ad un "estremo" della regione ammissibile (uno dei vertici del poligono).

Probabilmente, se avessimo scelto altri vettori entranti/uscenti nell'algoritmo avremmo trovato una soluzione ottimale diversa, ma saranno sempre corrispondenti ad un vertice del poligono evidenziato (ovvero della regione ammissibile).

Esempio 8.4.5 Vogliamo massimizzare la funzione $p(x_1, x_2) = x_1 + x_2$, con i vincoli (scritti direttamente con gli scarti per risparmiare tempo)

$$\begin{cases} x_1 + 2x_2 + x_3 = 10 \\ 2x_1 + x_2 + x_4 = 10 \\ x_1 + x_2 + x_5 = 6 \end{cases}$$

Ormai abbiamo capito come funziona, quindi scriviamo direttamente la tabella del simplesso:

				A_{v_1}	A_{v_2}	A_{v_3}	
		A_1	A_2	A_3	A_4	A_5	B
x_3	0	1	2	1	0	0	10
x_4	0	2	1	0	1	0	10
x_5	0	1	1	0	0	1	6
		-1	-1	0	0	0	0

Abbiamo degli indici con $z_s - c_s < 0$, stavolta con stesso valore assoluto. La scelta stavolta è totalmente arbitraria, quindi scegliamo come **vettore entrante** A_1 ($s = 1$).

$$\frac{\beta_1}{\alpha_{11}} = 10 \qquad \frac{\beta_2}{\alpha_{21}} = 5 \qquad \frac{\beta_3}{\alpha_{31}} = 6$$

il minimo è il secondo, quindi il **vettore uscente** sarà $A_{v_2} = A_4$.

$$\begin{array}{c} \begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 10 \\ 2 & 1 & 0 & 1 & 0 & 10 \\ 1 & 1 & 0 & 0 & 1 & 6 \end{pmatrix} \\ \begin{array}{cccccc} \uparrow & & \circlearrowleft & \downarrow & \circlearrowleft & \\ \textit{entra} & & \textit{resta} & \textit{esce} & \textit{resta} & \end{array} \end{array} \longrightarrow \begin{pmatrix} 0 & * & 1 & * & 0 & * \\ 1 & * & 0 & * & 0 & * \\ 0 & * & 0 & * & 1 & * \end{pmatrix}$$

Facendo i calcoli:

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 0 & 10 \\ 2 & 1 & 0 & 1 & 0 & 10 \\ 1 & 1 & 0 & 0 & 1 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & \frac{3}{2} & 1 & -\frac{1}{2} & 0 & 5 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 5 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 & 1 \end{pmatrix}$$

Riscriviamo dunque la tabella del simplesso:

		A_{v_2}		A_{v_1}		A_{v_3}	
		A_1	A_2	A_3	A_4	A_5	B
x_3	0	0	$\frac{3}{2}$	1	$-\frac{1}{2}$	0	5
x_1	1	1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	5
x_5	0	0	$\frac{1}{2}$	0	$-\frac{1}{2}$	1	1
		0	$-\frac{1}{2}$	0	$\frac{1}{2}$	0	5

Bisogna ripetere perché ci sono ancora degli indici con risultato negativo.

Vettore entrante A_2 ($s = 2$).

$$\frac{\beta_1}{\alpha_{12}} = \frac{10}{3} \qquad \frac{\beta_2}{\alpha_{22}} = 10 \qquad \frac{\beta_3}{\alpha_{32}} = 2$$

il minimo è il terzo, quindi il **vettore uscente** sarà $A_{v_3} = A_5$.

$$\begin{array}{cccccc} \multicolumn{6}{c}{\begin{pmatrix} 0 & \frac{3}{2} & 1 & -\frac{1}{2} & 0 & 5 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 5 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 & 1 \end{pmatrix}} \\ \circlearrowleft & \uparrow & \circlearrowleft & & \downarrow & \\ \textit{resta} & \textit{entra} & \textit{resta} & & \textit{esce} & \end{array} \longrightarrow \begin{pmatrix} 0 & 0 & 1 & * & * & * \\ 1 & 0 & 0 & * & * & * \\ 0 & 1 & 0 & * & * & * \end{pmatrix}$$

Facciamo i calcoli:

$$\begin{pmatrix} 0 & \frac{3}{2} & 1 & -\frac{1}{2} & 0 & 5 \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 5 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 2 \\ 1 & 0 & 0 & 1 & -1 & 4 \\ 0 & 1 & 0 & -1 & 2 & 2 \end{pmatrix}$$

E riscriviamo la tabella del simplesso:

		A_{v_2} A_1	A_{v_3} A_2	A_{v_1} A_3	A_4	A_5	B
x_3	0	0	0	1	1	0	2
x_1	1	1	0	0	1	-1	4
x_2	1	0	1	0	-1	2	2
		0	0	0	0	1	6

Siamo quindi arrivati ad una soluzione ottima:

$$\overline{x} = (4, 2, 2, 0, 0)$$

Capitolo 9

Considerazioni Finali

9.1 Retta tangente ad una curva

22/05/2014

La retta tangente ad una curva γ, scritta come $y = f(x)$, nel punto $(x_0, f(x_0))$ è esprimibile come

$$y = f(x_0) + m(x - x_0)$$
$$f(x) - \big(f(x_0) + m(x - x_0)\big) = o(x - x_0)$$

Prendiamo la curva γ descritta da $f(x, y) = 0$, e cerchiamo una retta tangente nel punto $P_0 = (x_0, y_0) \in \gamma$, che sarà del tipo $a(x - x_0) + b(y - y_0) = 0$. Da questa espressione ricaviamo:

$$f(x, y) = f(x_0, y_0) + \delta_x f(x_0, y_0)(x - x_0) + \delta_y f(x_0, y_0)(y - y_0) + o(||P - P_0||)$$
$$\implies f(x, y) - \big(\delta_x f(x_0, y_0)(x - x_0) + \delta_y f(x_0, y_0)(y - y_0)\big) = o(||P - P_0||)$$

E che succede se il gradiente è nullo? Risulta $f(x, y) = o(||P - P_0||)$, che è come dire "tutte le rette sono tangenti"... non ha molto senso.

Esempio 9.1.1 Data la curva γ: $y^2 - x^2 - x^3 = 0$, se cerchiamo la retta tangente al punto $(x_0, y_0) = (0, 0)$ risulta

$$grad\ f = (-2x - 3x^2, 2y) \Longrightarrow grad\ f(0, 0) = (0, 0)$$

quindi tutte le rette sono tangenti? No... Si ragiona così:

$$f(x, y) = f(x_0, y_0) + \delta_x f \cdot (x - x_0) + \delta_y f \cdot (y - y_0)+$$
$$+\frac{1}{2}\big(\delta_{xx} f \cdot (x - x_0)^2 + 2\delta_{xy} f \cdot (x - x_0)(y - y_0) + \delta_{yy} f \cdot (y - y_0)^2\big) - o(||P - P_0||^2) =$$

$$= f(x_0, y_0) + \big((x - x_0)\delta_x + (y - y_0)\delta_y\big) f(x_0, y_0) +$$
$$+\frac{1}{2}\big((x - x_0)\delta_x + (y - y_0)\delta_y\big)^2 f(x_0, y_0) - o(||P - P_0||) = \circledast$$

Così, in generale possiamo scrivere il *polinomio di Taylor* di una funzione in più variabili nella forma

$$T_n(x, y) = f(x_0, y_0) + \sum_{k=1}^{n} \frac{\big((x - x_0)\delta_x + (y - y_0)\delta_y\big)^k}{k!} f(x_0, y_0)$$

Torniamo alla retta tangente: dopo tutti i calcoli fatti scriviamo

$$f(x, y) = \frac{1}{2}\delta_{xx} f \cdot (x - x_0)^2 + \delta_{xy} f \cdot (x - x_0)(y - y_0) + \frac{1}{2}\delta_{yy} f \cdot (y - y_0)^2 + o(||P - P_0||)^2$$

e ci chiediamo se esistono $a, b, c \in \mathbb{R}$ tali che

$$f(x, y) - \big(a(x - x_0)^2 + b(x - x_0)(y - y_0) + c(y - y_0)^2\big) = o(||P - P_0||)^2$$

La risposta è che esistono e sono unici: sono proprio quelli scritti sopra!

$$\begin{cases} a = \frac{1}{2}\delta_{xx} f \\ b = \delta_{xy} f \\ c = \frac{1}{2}\delta_{yy} f \end{cases}$$

Quindi per la curva data $y^2 - x^2 - x^3 = 0$, nel punto $P_0 = (0, 0)$ otteniamo

$$grad\ f = (-2x - 3x^2, 2y)$$

$$\delta_{xx} f = -2 \qquad \delta_{xy} f = 0 \qquad \delta_{yy} f = 2$$

per cui esistono <u>due</u> rette tangenti, di equazione complessiva

$$-\frac{1}{2}(-2x^2 + 2y^2) = 0 \qquad\qquad \text{con } y^2 - x^2 = 0$$

Osservazione Un polinomio omogeneo rappresenta un'insieme di rette. Vediamone un esempio: $r(x', y') = 3x'^2 - 5x'y' - 8y'^2$.

Se $r(x_0, y_0) = 0$, allora $r(tx_0, ty_0) = t^2 r(x_0, y_0) = 0$.

Quindi il luogo dei punti che verificano $r(x, y) = 0$ rappresenta proprio una retta passante per l'origine!

9.2 Classificazione delle Coniche

Una *Conica* è una curva definita da un'equazione di questo tipo:

$$f(x,y) = a_{11}x^2 + 2a_{12}xy + 2a_{13}x + a_{22}y^2 + 2a_{23}y + a_{33} = 0$$

Il *discriminante* della conica γ di equazione $f(x,y) = 0$ è la matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

E definiamo

$$B := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad det\ B := A_{33}$$

Se $detA = 0$, γ si dice *degenere*.

Teorema 9.2.1 *Una conica degenere è unione di due rette.*

Dimostrazione. Ci sono diversi casi da analizzare:

1. $\boxed{detB \neq 0}$. Il sistema

$$\begin{cases} a_{11}x + a_{12}y + a_{12} = 0 \\ a_{12}x + a_{22}y + a_{23} = 0 \end{cases}$$

è di Cramer, ed ha una sola soluzione $(x_0, y_0) = P_0 \in \gamma$.

La forma quadratica $f(x,y)$ si può scrivere equivalentemente come

$$f(x,y) = \big(x, y, 1\big) \cdot A \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \left\langle \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, A \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \right\rangle$$

Scriviamo il polinomio di Taylor[1] di f:

$$f(x,y) = \frac{1}{2}\big((x - x_0)\delta_x + (y - y_0)\delta_y\big)^2 f(x_0, y_0)$$

Se calcoliamo le derivate seconde

$$\delta_{xx}f = 2a_{11} \qquad \delta_{xy}f = 2a_{12} \qquad \delta_{yy}f = 2a_{22}$$

[1]Tenendo conto del fatto che il polinomio di Taylor approssima una qualsiasi funzione in forma polinomiale, quindi se lo scriviamo di un polinomio possiamo anche ottenere il polinomio stesso, se sviluppiamo Taylor con abbastanza gradi!

l'equazione di γ è dunque:

$$a_{11}(x-x_0)^2 + 2a_{12}(x-x_0)(y-y_0) + a_{22}(y-y_0)^2 = 0$$

$$\Longrightarrow \frac{\Delta}{4} = a_{12}^2 - a_{11}a_{22} = -A_{33}$$

- $A_{33} < 0 \Longrightarrow \gamma =$ due rette reali incidenti
  Esempio: $x^2 - y^2 = 0 \Longrightarrow (x-y)(x+y) = 0$

$$\Longrightarrow A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- $A_{33} > 0 \Longrightarrow \gamma =$ due rette complesse coniugate
  Esempio: $x^2 + y^2 = 0 \Longrightarrow (x-iy)(x+iy) = 0$

$$\Longrightarrow A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

2. $\boxed{detB = 0}$.

- Supponiamo inizialmente $a_{11} \neq 0$, per cui riscriviamo la matrice A come

$$A = \begin{pmatrix} 1 & a & b \\ a & a^2 & c \\ b & c & d \end{pmatrix}$$

Visto che la conica è degenere si dovrà avere $detA = 0$, ovvero

$$detA = a^2d + 2abc - a^2b^2 - c^2 - a^2d = -(c-ab)^2$$

$$\Longrightarrow \quad c = ab$$

Abbiamo trovato $c = ab$, quindi riscriviamo

$$A = \begin{pmatrix} 1 & a & b \\ a & a^2 & ab \\ b & ab & d \end{pmatrix}$$

e riscriviamo anche la curva γ:

$$x^2 + 2axy + 2bx + a^2y^2 + 2aby + d = 0$$

Raccogliendo in modo furbo:

$$x^2 + 2axy + 2bx + a^2y^2 + 2aby + d + b^2 = b^2$$

$$\implies \quad (x + ay + b)^2 = b^2 - d$$

Quindi, se $b^2 - d > 0$, γ rappresenta due rette parallele in $\mathbb{R}$.

- Supponiamo adesso che $a_{11} = 0$, per cui

$$A = \begin{pmatrix} 0 & 0 & \beta \\ 0 & \alpha & \delta \\ \beta & \delta & \epsilon \end{pmatrix}$$

Visto che $\alpha \neq 0$, deve essere $\beta = 0$, quindi

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha & \delta \\ 0 & \delta & \epsilon \end{pmatrix}$$

$$\implies \quad \alpha y^2 + 2\delta y + \epsilon = 0$$

quindi in questo caso γ può rappresentare tutti i casi al variare di α, δ, ϵ dell'equazione.

□

Passiamo adesso alle coniche non degeneri, ovvero con $detA \neq 0$.

Coniche a centro: la conica si dice "a centro" se $detA \neq 0$, $detB \neq 0$.

Visto che $detB \neq 0$, possiamo fare un ragionamento simile a prima:

$$\begin{cases} a_{11}x + a_{12}y + a_{12} = 0 \\ a_{12}x + a_{22}y + a_{23} = 0 \end{cases}$$

è di Cramer, ed ha una sola soluzione $(x_0, y_0) = P_0$. Stavolta $P_0 \notin \gamma$ perché il punto non può soddisfare anche la terza equazione $a_{13}x + a_{23}y + a_{33} = 0$.

Troviamo lo sviluppo di Taylor:

$$f(x, y) = f(x_0, y_0) + a_{11}(x - x_0)^2 + 2a_{12}(x - x_0)(y - y_0) + a_{22}(y - y_0)^2$$

Allora P_0 è centro di simmetria per γ[2].

[2]Proprio per questo si chiamano coniche "a centro"!

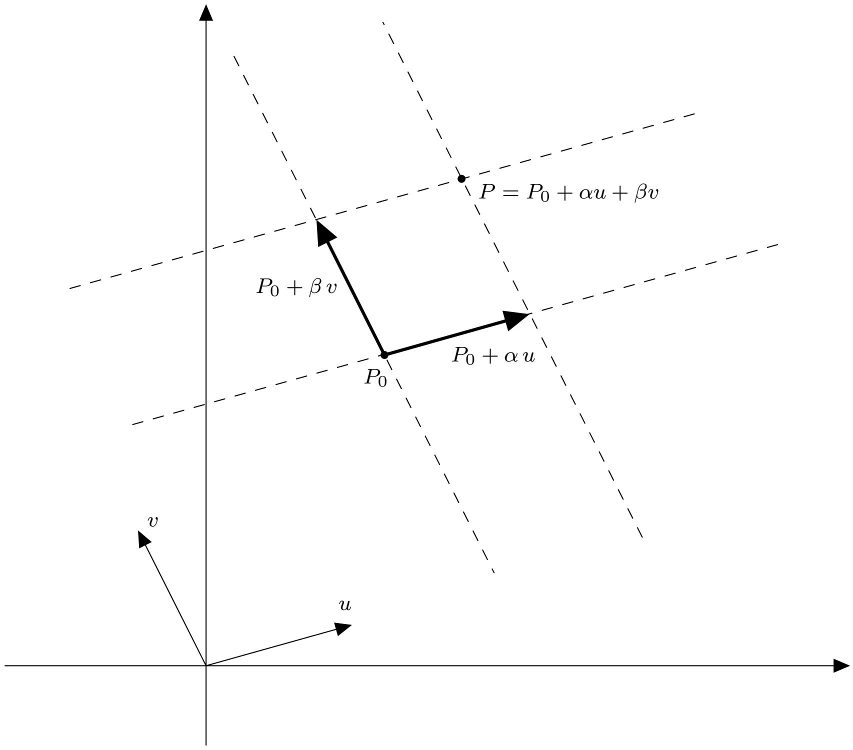
Infatti, presi due punti simmetrici rispetto a P_0, che chiamiamo $S = (x_0 + h, y_0 + k)$ e $Q = (x_0 - h, y_0 - k)$, risulta

$$f(Q) = f(x_0, y_0) + a_{11}h^2 + 2a_{12}hk + a_{22}k^2 = f(S)$$

quindi $Q \in \gamma \Longleftrightarrow S \in \gamma$.

B ha due autovalori reali non nulli (poiché $detB \neq 0$) λ, μ, con autovettori corrispondenti u, v. Essendo B simmetrica, questi due devono essere ortogonali[3].

Nulla ci vieta di prendere gli autovettori di norma unitaria, per facilitare i conti. La situazione è schematizzata con il seguente disegno:



Possiamo trovare il valore di f nel punto $P = P_0 + \alpha u + \beta v$ calcolando

$$f(P) = f(P_0) + < P - P_0, B(P - P_0) > = f(P_0) + < \alpha u + \beta v, B(\alpha u + \beta v) > =$$

sappiamo che $Bu = \lambda u$, $Bv = \mu v$, $< u, v > = 0$, quindi

$$= f(P_0) + < \alpha u + \beta v, \lambda\alpha u + \mu\beta v > = f(P_0) + \lambda\alpha^2 + \mu\beta^2$$

[3]Si dimostra molto facilmente tramite una serie di uguaglianze: $\mu < u, v > = < u, Bv > = < Bu, v > = \lambda < u, v >$

L'equazione di γ nel sistema di riferimento (α, β) è

$$f(P_0) + \lambda\alpha^2 + \mu\beta^2 = 0$$

questa è di nuovo l'equazione della conica, anche se scritta in un sistema di riferimento diverso, da cui possiamo riscrivere

$$A' = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & f(P_0) \end{pmatrix}$$

quindi al variare dei segni di λ, μ e $f(P_0)$ otteniamo ellissi, iperboli o parabole.