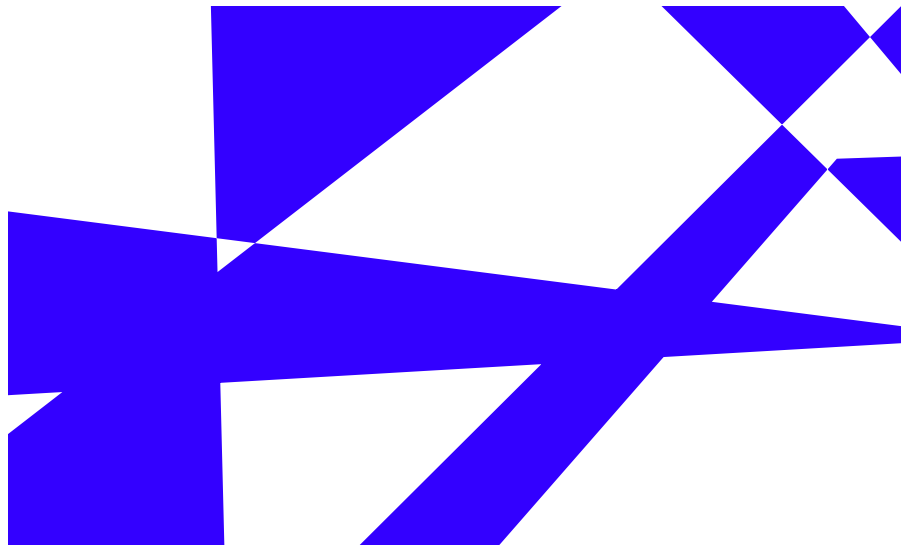


# La convessità in

- . . . ottica
- . . . **geometria**
- . . . **analisi**
- . . . finanza



**Simon Salamon**

<http://calvino.polito.it/~salamon>

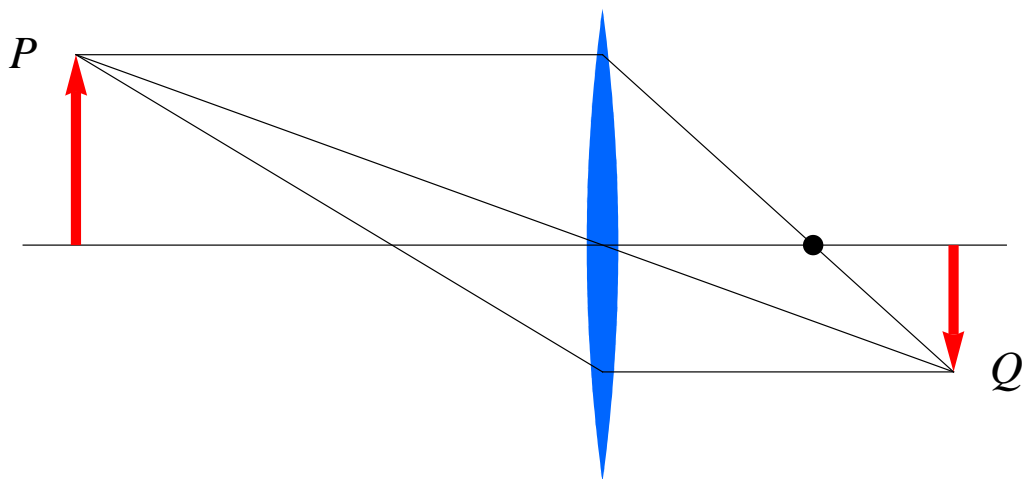
# Indice

1. L'equazione di una lente
2. Convessità descrittiva
3. Vettori e combinazioni convesse
4. Funzioni convesse
5. Disuguaglianze e norme
6. Il prezzo di un'obbligazione

## 1. L'equazione di una lente

La corrispondenza tra un oggetto  $P$  e la sua immagine  $Q$  tramite una lente convessa è simbolica di un fenomeno di dualità che esiste in altri campi della fisica e anche nella matematica.

Per semplicità, sia la distanza focale  $f=1$ . Raggi da  $P$  ortogonali alla lente  $\ell$  passano per il punto focale indicato. Supponiamo che  $P$  abbia altezza  $a$  e sia a distanza  $p$  da  $\ell$ , mentre l'immagine  $Q$  altezza  $b$  e distanza  $q$  da  $\ell$ .



Allora

$$\frac{a}{p-1} = \frac{b}{1}, \quad \frac{a}{1} = \frac{b}{q-1},$$

da cui

$$(p-1)(q-1) = 1 \Rightarrow \frac{1}{p} + \frac{1}{q} = 1.$$

Un caso importante è quello in cui  $p = q = 2$ : un oggetto messo a due volte la distanza focale avrà immagine la stessa distanza dall'altra parte.

## 2. Convessità descrittiva

L'intervallo chiuso definito da  $a, b \in \mathbb{R}$  viene scritto  $[a, b]$ . Dati i punti  $P, Q \in \mathbb{R}^n$ , è quindi naturale indicare con  $[P, Q]$  il segmento chiuso (della retta) tra  $P$  e  $Q$ . (Meglio supporre che  $n = 2$  o  $3$  per ora, perchè la definizione di segmento in  $\mathbb{R}^n$  viene fatta in §3 in termini vettoriali.) Un insieme è convesso se contiene tale segmento per ciascuna coppia  $P, Q$  dei suoi punti:

Definizione Sia  $\Omega \subseteq \mathbb{R}^n$ . Allora  $\Omega$  è convesso se

$$P, Q \in \Omega \Rightarrow [P, Q] \in \Omega.$$

Esercizio  $\Omega \subseteq \mathbb{R}$  è convesso se e solo se  $\Omega$  è un intervallo, cioè  $(a, b), [a, b), (a, b], [a, b]$ , eventualmente con  $a$  o  $b = \infty$ .

La convessità di  $\Omega \subset \mathbb{R}^n$  non dipende dall'inclusione dei punti del 'bordo'. Tecnicamente, l'unione di  $\Omega$  con i suoi 'punti limite' si chiama la chiusura  $\overline{\Omega}$ ; è noto che se  $\Omega$  è convesso anche  $\overline{\Omega}$  lo è.

L'intersezione di due (o qualsiasi collezione di) insiemi convessi è sempre convessa. Questa stessa proprietà vale (ad esempio) per sottospazi di uno spazio vettoriale. Possiamo quindi dare la

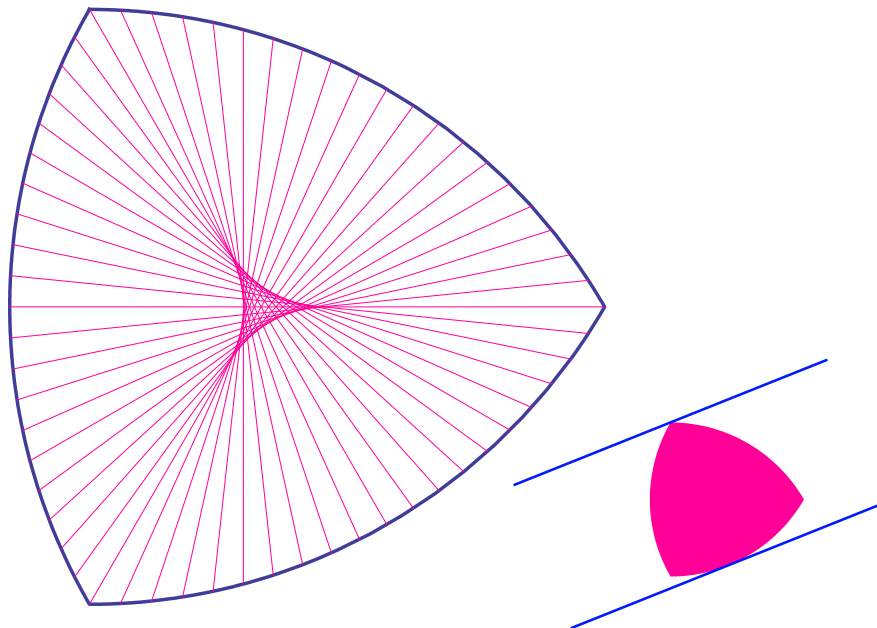
Definizione Sia  $S \subset \mathbb{R}^n$ . L'involuppo convesso (detto anche convex hull)  $C(S)$  di  $S$  è l'intersezione di tutti gli insiemi convessi di  $\mathbb{R}^n$  che contengono  $S$ .

Segue che  $C(S)$  è l'insieme convesso più piccolo che contiene  $S$ : è l'unico insieme convesso che soddisfa

$$S \subseteq T, \quad T \text{ convesso} \Rightarrow C(S) \subseteq T.$$

Qualsiasi triangolo è il bordo di una regione convessa, essendo l'intersezione di 3 semipiani. Anche l'interno di una circonferenza è una regione convessa limitata. Si può costruire un esempio un po' meno familiare combinando questi due oggetti.

Prendiamo un triangolo equilatero e facciamolo 'ingrassare' rimpiazzando i lati con archi di cerchi. Prendendo i tre vertici come centri delle circonferenze, si ottiene il 'triangolo' di Reuleaux, un esempio di una 'monetina' di larghezza costante, usato anche per trapanare dei buchi quadrati (come l'ultimo a pagina 13).



Se i tre vertici stanno su una circonferenza unitaria, la larghezza della monetina è  $\sqrt{3}$ . (La presenza di  $\sqrt{3}$  spesso segnala una simmetria di ordine 3.)

Esercizio Verificare che la lunghezza del perimetro è  $\pi\sqrt{3}$ , quella di una circonferenza di diametro  $\sqrt{3}$ . Vale lo stesso risultato per qualsiasi curva di larghezza  $\sqrt{3}$  costante. Le rette indicate collegano punti sul bordo del 'triangolo' a metà distanza  $\pi\sqrt{3}/2$ .

### 3. Vettori e combinazioni convesse

Rappresentiamo due punti  $P_1, P_2$  nel piano con vettori del tipo  $\mathbf{v}_1 = (x_1, y_1)$  e  $\mathbf{v}_2 = (x_2, y_2)$ . Lo spostamento da  $\mathbf{v}_2$  a  $\mathbf{v}_1$  corrisponde al vettore

$$\mathbf{v}_1 - \mathbf{v}_2 = (x_1 - x_2, y_1 - y_2).$$

Quindi un punto  $P$  del segmento  $[P_1, P_2]$  è dato da

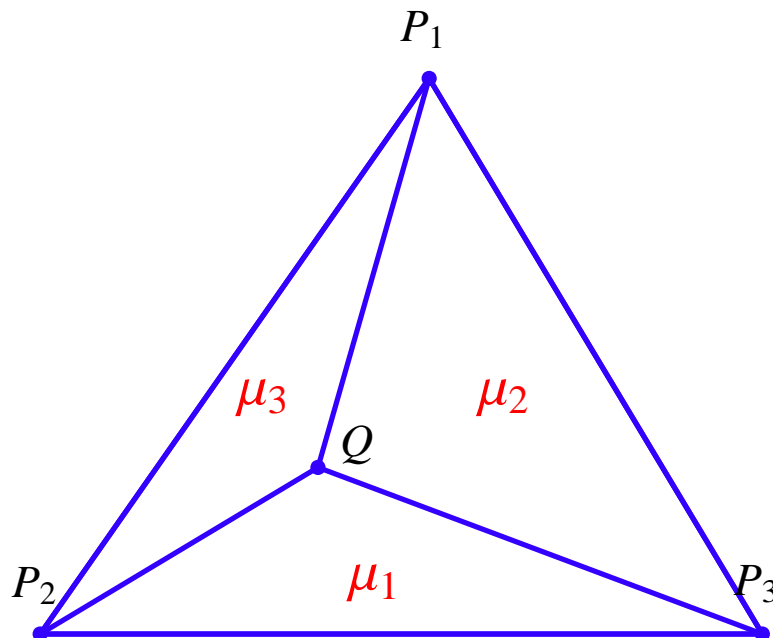
$$\begin{aligned}\mathbf{v} &= \mathbf{v}_1 + \lambda(\mathbf{v}_1 - \mathbf{v}_2) = \lambda\mathbf{v}_1 + (1 - \lambda)\mathbf{v}_2 \\ &= \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2,\end{aligned}$$

dove  $\lambda_i \geq 0$  e  $\lambda_1 + \lambda_2 = 1$ . Osservare che  $P$  divide  $P_1P_2$  nel rapporto  $\lambda_2 : \lambda_1$ ; se  $|P_1P_2|$  fosse 1 allora avremmo  $|P_1P| = \lambda_2$  e  $|PP_2| = \lambda_1$ .

Adesso consideriamo tre punti, vertici di un triangolo. Un punto  $Q$  interno avrà vettore

$$\begin{aligned}\mu\mathbf{v} + (1 - \mu)\mathbf{v}_3 &= \mu(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2) + (1 - \mu)\mathbf{v}_3 \\ &= \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \mu_3\mathbf{v}_3,\end{aligned}$$

dove  $\mu_1 + \mu_2 + \mu_3 = 1$ .



Si può dimostrare che, questa volta, il coefficiente  $\mu_i$  che aiuta a determinare  $Q$  corrisponde (in proporzione) all'area di quel triangolo che non tocca  $P_i$ .

La seguente generalizzazione ci dà una rappresentazione analitica dell'involuppo convesso di un insieme finito.

Lemma Sia  $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset \mathbb{R}^N$ . Allora

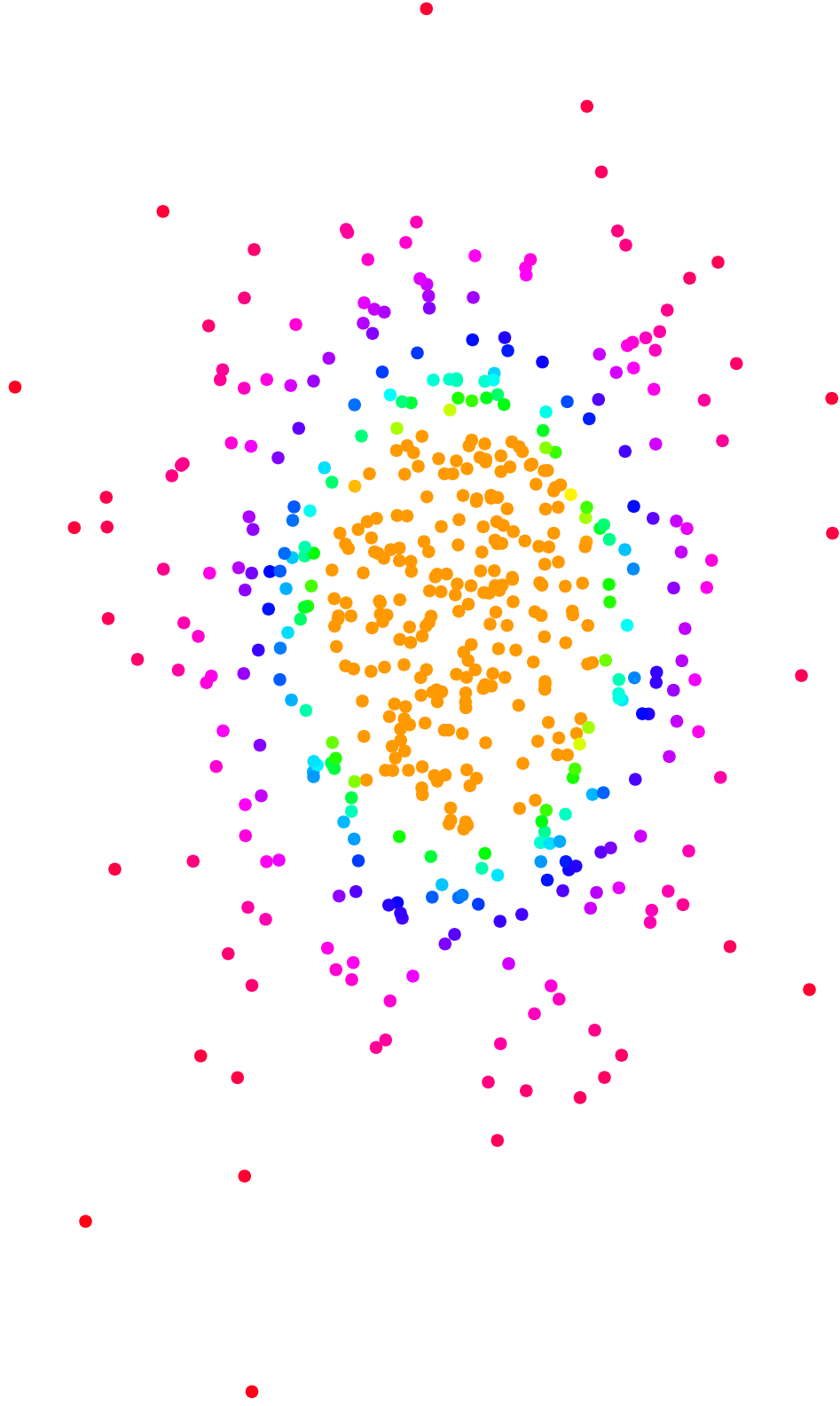
$$C(S) = \left\{ \sum_{i=1}^n \lambda_i \mathbf{v}_i : \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0 \right\}.$$

Dimostrazione: Sia  $T$  un insieme convesso che contiene tutti i punti  $\mathbf{v}_i$ . L'osservazione cruciale è che

$$\sum_{i=1}^n \lambda_i \mathbf{v}_i = \lambda_1 \mathbf{v}_1 + (1 - \lambda_1) \sum_{i=2}^n \lambda'_i \mathbf{v}_i$$

dove  $\lambda'_2 + \dots + \lambda'_n = 1$ . (Questo segue perchè  $1 - \lambda_1 = \lambda_2 + \dots + \lambda_n$ .) Dimostrando il lemma per induzione in  $n$ , potremmo supporre che la seconda somma stia già in  $T$ . Risulta che  $C(S) \subseteq T$ . Per concludere, basta verificare che  $C(S)$  stesso è convesso, un esercizio facile. QED

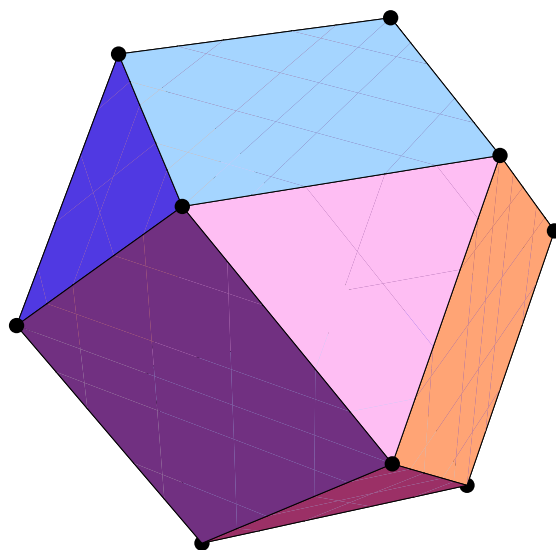
È molto studiato il problema di trovare meccanicamente l'involuppo convesso di un insieme di (tanti) punti nel piano. Basta elencare i 'punti estremi', quelli determinati da una corda elastica messa intorno alla figura (più rosso nell'esempio seguente). Ci sono diversi algoritmi in uso; quelli più veloci evitano l'obbligo di analizzare tutti i punti in successione, una alla volta.



Si ottengono dei bei involucri nello spazio scegliendo un piccolo insieme di punti con un'opportuna simmetria. I 6 punti

$$(\pm 1, 0, 0), \quad (0, \pm 1, 0), \quad (0, 0, \pm 1)$$

definiscono l'ottaedro e gli 8 punti  $(\pm 1, \pm 1, \pm 1)$  sono i vertici di un cubo. Gli altri solidi platonici (tetraedro, icosaedro, dodecaedro) si trovano in modo analogo.



Questo solido è

$$C(S) = \left\{ \sum_{i=1}^{12} \lambda_i \mathbf{v}_i : \sum_{i=1}^{12} \lambda_i = 1, \quad \lambda_i \geq 0 \right\},$$

in cui i 12 punti si ottengono da  $\mathbf{v}_1 = (1, 1, 0)$  permutando le coordinate e cambiando segni. È un cubo con gli 8 vertici 'tagliati'; risultano quindi 8 triangoli e 6 quadrati (uno dentro ogni faccia).

Esercizio Verificare che i punti  $(0, 0, \sqrt{2})$ ,  $(1, 0, 0)$ ,  $(-\frac{1}{2}, \frac{\sqrt{3}}{2}, 0)$ ,  $(-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 0)$  sono i vertici di un tetraedro regolare.

## 4. Funzioni convesse

Studieremo una classe di funzioni  $f: (a, b) \rightarrow \mathbb{R}$ .

Definizione  $f$  è convesso se

$$\lambda_1 + \lambda_2 = 1 \Rightarrow f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

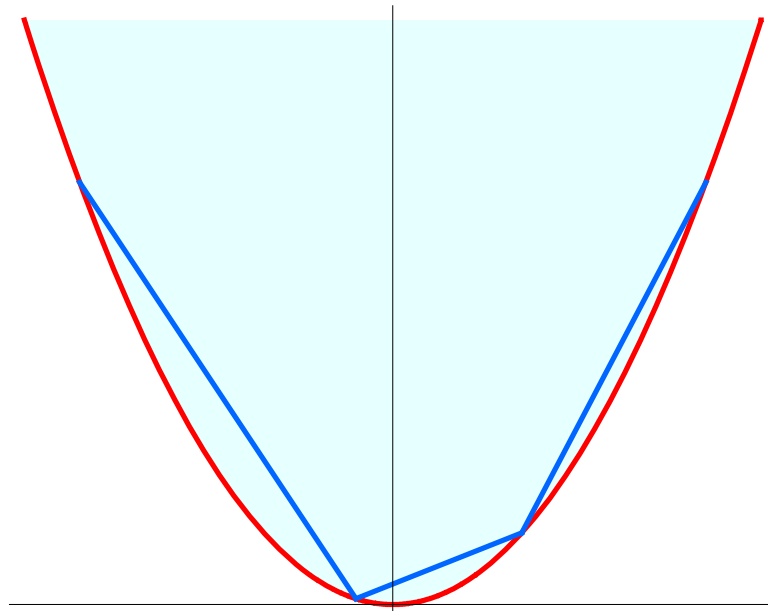
Sia  $\mathbf{v}_i = (x_i, f(x_i))$ . Sappiamo che

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = (\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 f(x_1) + \lambda_2 f(x_2))$$

è un punto sul segmento tra  $\mathbf{v}_1$  e  $\mathbf{v}_2$ . La definizione vuol dire che il grafico di  $f$  tra  $\mathbf{v}_1$  e  $\mathbf{v}_2$  non sta mai sopra il segmento associato. Equivale a dire che la regione

$$\{(x, y) \in \mathbb{R}^2 : f(x) \leq y\}$$

verticalmente sopra il grafico è convesso.



Se  $-f$  è convessa, diciamo che  $f$  è concava. Ogni minimo relativo di una funzione convessa deve essere un minimo assoluto.

È possibile dimostrare che ogni funzione convessa su  $(a, b)$  è automaticamente continua, cioè il grafico non ha 'buchi'. Se  $f$  è convessa allora

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a) + f(b)}{2}.$$

Viceversa, questa condizione basta per concludere che  $f$  sia convessa se sappiamo già che  $f$  è continua.

Siano  $a < b < c$ . Sia  $f$  una funzione convessa definita su un intervallo che contiene  $[a, c]$ . Allora

$$b = \lambda_1 a + \lambda_2 c = \frac{c-b}{c-a}a + \frac{b-a}{c-a}c$$

$$\Rightarrow f(b) \leq \frac{c-b}{c-a}f(a) + \frac{b-a}{c-a}f(c).$$

Ricordando che  $\lambda_1 + \lambda_2 = 1$ ; sottraendo  $f(b)$ , abbiamo

$$0 \leq \frac{c-b}{c-a}(f(a) - f(b)) + \frac{b-a}{c-a}(f(c) - f(b))$$

$$\Rightarrow \frac{f(b) - f(a)}{b-a} \leq \frac{f(c) - f(b)}{c-b}.$$

Vuol dire che la pendenza di segmenti adiacenti cresce andando a destra. Se  $b$  è fissato e  $a \rightarrow b$  allora il termine a sinistra tende a  $f'(b)$ . Adesso si scelgano tre segmenti definiti da  $a < b < c < d$ , si fissi quello in mezzo, e si ponga  $a \rightarrow b$  e  $d \rightarrow c$ . Così  $f'(b) \leq f'(c)$ .

Teorema Se  $f$  è differenziabile,  $f$  è convessa se e solo se  $f'$  è crescente. Se  $f$  è due volte differenziabile,  $f$  è convessa se e solo se  $f''$  è sempre non-negativa.

Esempi Le seguenti funzioni sono convesse:

$$\begin{aligned} & x^\alpha \text{ per } \alpha \leq 0 \text{ e } \alpha \geq 1 \text{ su } (0, \infty), \\ & -x^\alpha \text{ per } 0 \leq \alpha \leq 1 \text{ su } [0, \infty), \\ & e^x \text{ su } \mathbb{R}, \quad -\log x \text{ su } (0, \infty), \\ & -\sin x \text{ su } [0, \pi]. \end{aligned}$$

Sia  $f: (a, b) \rightarrow \mathbb{R}$  una funzione convessa.

Lemma: disuguaglianza di Jensen

$$(J) \quad \text{Se } \sum_{i=1}^n \lambda_i = 1 \text{ allora } f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Dimostrazione: Per induzione in  $n$ , usando

$$\sum_{i=1}^n \lambda_i x_i = \lambda_1 x_1 + (1 - \lambda_1) \sum_{i=2}^n \lambda'_i x_i, \quad \sum_{i=2}^n \lambda'_i = 1,$$

come nella costruzione di  $C(S)$ . QED

In particolare,

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n}.$$

Esercizi. 1 Applicare l'ultima equazione con  $f(x) = e^x$  per dimostrare la disuguaglianza delle medie

$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n}.$$

2. Sia  $x_i \geq 0$  e si scriva  $N_p = \left(\frac{x_1^p + \cdots + x_n^p}{n}\right)^{1/p}$ .

Allora  $p < q \Rightarrow N_p \leq N_q$ .

## 5. Disuguaglianze e norme

La lunghezza di un vettore  $\mathbf{x} = (x_1, x_2, x_3)$  in  $\mathbb{R}^3$  è

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2} = \left( \sum_{i=1}^3 x_i^2 \right)^{1/2}.$$

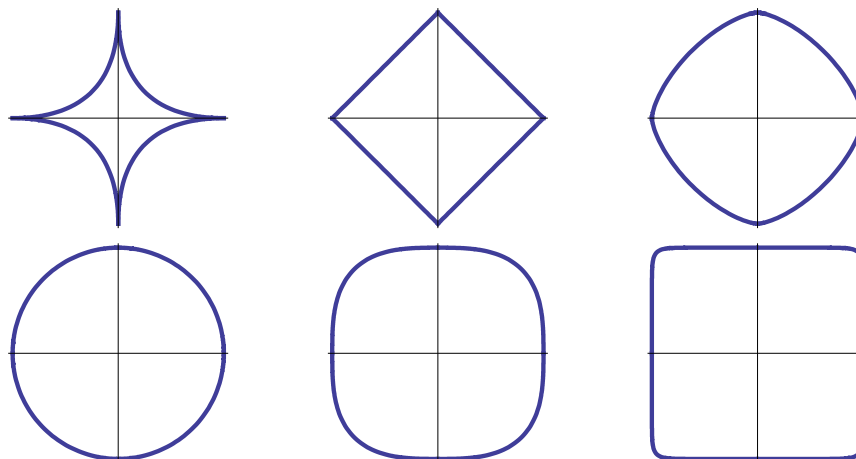
Si chiama la norma di  $\mathbf{x}$ . Se  $P, Q$  sono punti nello spazio rappresentati da  $\mathbf{x}, \mathbf{y}$  allora  $\|\mathbf{x} - \mathbf{y}\| = |PQ|$  è la distanza tra  $P$  e  $Q$ .

In realtà, possiamo definire una quantità analoga

$$\|\mathbf{x}\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}$$

per  $\mathbf{x} \in \mathbb{R}^n$  e  $p > 0$ . La scelta dalle potenze reciproche garantisce che  $\|\lambda \mathbf{v}\|_p = \lambda \|\mathbf{v}\|_p$  se  $\lambda > 0$ .

Il caso sopra corrisponde a  $p = 2$ , ma per altri valori di  $p \neq 2$  otteniamo una 'distanza' distorta. Nel piano, cioè per  $n = 2$ , i disegni che seguono rappresentano i vettori per cui la ' $p$ -norma' (e quindi la 'distanza' dall'origine) è uguale a 1 per  $p = \frac{1}{2}, 1, \frac{3}{2}$  (sopra) e  $p = 2, 3, 10$  (sotto).



Siano  $\mathbf{x} = (x_1, \dots, x_n)$  e  $\mathbf{y} = (y_1, \dots, y_n)$  vettori in  $\mathbb{R}^n$ .  
(Per semplicità si può supporre che  $n=2$  o  $3$ ).

Teorema: le disuguaglianze di Minkowski e di Hölder

(M) Se  $p > 1$  allora  $\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$ .

(H) Se  $p, q > 1$  e  $\frac{1}{p} + \frac{1}{q} = 1$  allora  $\sum_{i=1}^n |x_i y_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q$ .

La prima è chiamata anche la disuguaglianza triangolare; è equivalente a

$$\|\mathbf{x} - \mathbf{y}\|_p \leq \|\mathbf{x} - \mathbf{z}\|_p + \|\mathbf{z} - \mathbf{y}\|_p.$$

Quando  $p = 2$ , dice che in un triangolo ogni lato è minore della somma degli altri due: camminando lungo una retta minimizza la distanza. Sempre con  $p = 2$ , (H) implica la disuguaglianza di Cauchy-Schwartz, che in  $\mathbb{R}^3$  è affermazione che  $|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ , evidente se il prodotto scalare viene definito in termine di  $\cos \theta$ .

Torniamo al caso generale. L'esercizio 5 sul foglio spiega che (M) segue da (H), che dimostriamo qui.

Lemma: la disuguaglianza di Young

(Y) Se  $a, b > 0$  e  $\frac{1}{p} + \frac{1}{q} = 1$  allora  $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ .

Dimostrazione di (Y): Siano  $\lambda_1 = \frac{1}{p}$ ,  $\lambda_2 = \frac{1}{q}$ . Dal fatto che  $\exp: x \mapsto e^x$  è convesso, abbiamo

$$\exp\left(\frac{x}{p} + \frac{y}{q}\right) \leq \frac{1}{p} e^x + \frac{1}{q} e^y$$

per ogni  $x, y \in \mathbb{R}$ . Si sostituisca  $e^x = a^p$ ,  $e^y = b^q$ . QED

Dimostrazione di (H): Siano  $a = |x_i|/\|\mathbf{x}\|_p$  e  $b = |y_i|/\|\mathbf{y}\|_q$  e applicare (Y). Il primo termine a destra di (Y) è  $1/p$  moltiplicato per  $a^p = |x_i|^p/\sum|x_i|^p$ . Sommando per  $i = 1, \dots, n$ , otteniamo

$$\frac{1}{\|\mathbf{x}\|_p\|\mathbf{y}\|_q} \sum_{i=1}^n |x_i||y_i| \leq \frac{1}{p} + \frac{1}{q} = 1,$$

e (H) segue. QED

Infine, si consideri

$$\mathcal{D}_p = \{\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\|_p \leq 1\},$$

il 'disco' unitario in  $\mathbb{R}^n$  (relativo alla  $p$ -norma). Il bordo di  $\mathcal{D}_p$  in  $\mathbb{R}^2$  è disegnato a pagina 13 per 6 valori di  $p$ . Mentre  $\mathcal{D}_2$  è il solito disco circolare,  $\mathcal{D}_1$  è un diamante.

Siano  $\mathbf{x}, \mathbf{y} \in \mathcal{D}_p$ . Se  $p > 1$ , (M) implica che

$$\begin{aligned} \|\lambda\mathbf{x} + (1-\lambda)\mathbf{y}\|_p &\leq \lambda\|\mathbf{x}\|_p + (1-\lambda)\|\mathbf{y}\|_p \\ &\leq \lambda + (1-\lambda) = 1, \end{aligned}$$

confermando che  $\mathcal{D}_p$  è convessa. È falsa per  $p < 1$ .

Esercizi 1. Spiegare perchè (M) vale anche per  $p = 1$ .

2. Verificare che (H) vale per  $p=1$  se prendiamo  $q=\infty$  e interpretiamo  $\|\mathbf{y}\|_\infty$  come il massimo dei numeri  $|y_i|$  per  $1 \leq i \leq n$ .

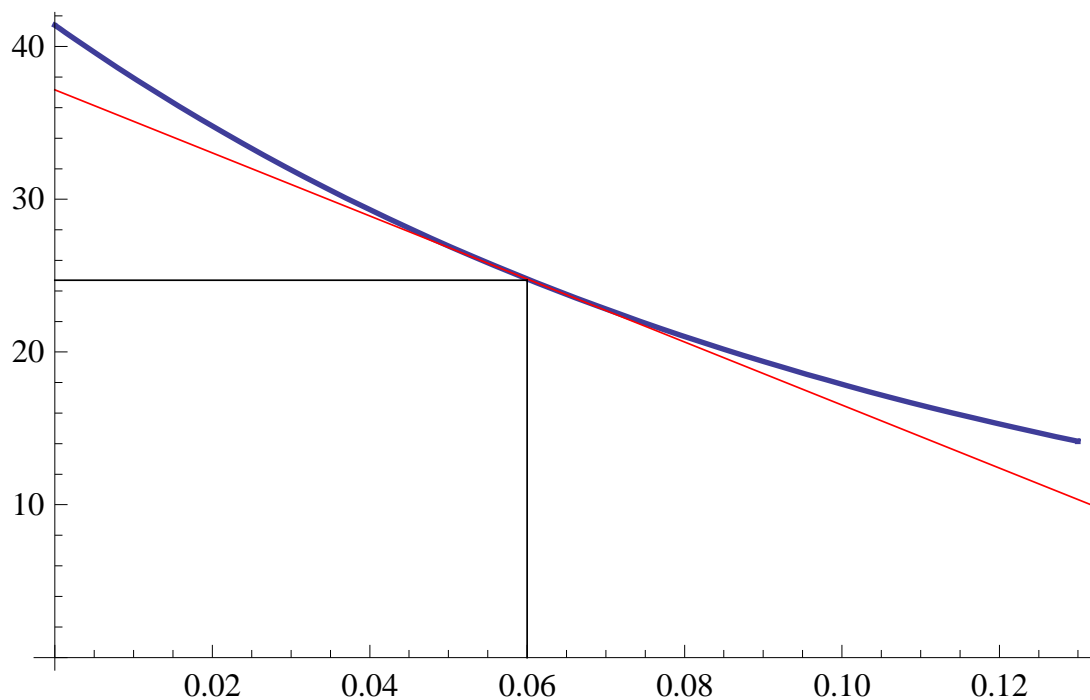
## 6. Il prezzo di un'obbligazione



Ricordate (MAT+ del 22/10) che una capitale  $F$  oggi dovrebbe valere

$$F \left(1 + \frac{r}{360}\right)^{360n} \approx Fe^{rn}$$

dopo  $n$  anni, con l'ipotesi di un tasso di interesse  $r$  costante, pagato ogni giorno. Viceversa, un pagamento  $F$  da fare tra  $n$  anni vale solo  $Fe^{-rn}$  oggi.



Il grafico indica il valore attuale

$$V(r) = \sum_{k=1}^{10} cFe^{-rk} + Fe^{-rn}$$

di un'obbligazione con pagamento finale di euro  $F = 30 \times 10^6$  dopo  $n = 10$  anni insieme a cedole annuali di  $cF = \frac{3.825}{100}F$ . È la somma che si dovrebbe teoricamente investire oggi per coprire le cedole e il pagamento finale, se fosse stabilito un tasso di interesse  $r$  costante durante il periodo.

Siccome

$$V''(r) = k^2 \sum_{k=1}^{10} cF e^{-rk} + n^2 F e^{-rn} > 0,$$

la funzione  $V(r)$  è convessa, anche se in un intervallo di  $r$  più realistico il grafico è ben approssimato dalla retta tangente.

Conclusione Mentre un aumento  $r \rightsquigarrow r + \Delta r$  del tasso abbassa il valore dell'obbligazione, la differenza  $|\Delta V|$  è meno di quello che risulta quando  $r \rightsquigarrow r - \Delta r$ .

Tali considerazioni sono importanti nell'acquisto di una selezione di obbligazioni con lo scopo di minimizzare ('hedge') i rischi associati con le incertezze del vero tasso di interesse\*. In realtà, ci serve la funzione inversa  $r = r(V)$  che esprime  $r$  in termini del prezzo  $V$ . Visto in questo senso, il tasso  $r$  è virtuale; misura invece il rendimento alla scadenza (o 'yield to maturity' YTM) dell'obbligazione sotto considerazione.

Avendo un ampio spettro di obbligazioni citate giorno per giorno sui mercati con diverse caratteristiche, si può anche costruire un grafico del rendimento  $r$  come funzione della durata (un vero 'bond curve') per ottenere una previsione del futuro tasso di interesse, che non sarà certamente costante.

Esercizio Sia  $D(r) = -V'(r)/V(r)$ . Verificare che una obbligazione con  $c=0$  ('zero-coupon bond') ha  $D(r)$  costante e uguale a  $n$ . In generale,  $D(r)$  misura la scadenza effettiva dell'obbligazione.

\*pagina 16