Appunti dalle lezioni del corso di

Teoria dei segnali

DISTRIBUZIONE GAUSSIANA UNIDIMENSIONALE

Prof.Alessandro Neri

INDICE :

2 - DISTRIBUZIONE GAUSSIANA UNIDIMENSIONALE

DEF.1- Una variabile aleatoria unidimensionale si dice gaussiana o a distribuzione normale, se la sua f.d.p. è del tipo:

$$P_x(x) = \frac{1}{\sqrt{2p}\, s} e^{-\frac{(x-m)^2}{2s^2}} \qquad (1)$$

con s^2 0.

Si noti che tale f.d.p. è completamente descritta da due parametri m e s^2 che rappresentano rispettivamente il valore atteso e la varianza della v.a. X. Per la verifica di cio' si può fare inizialmente riferimento alla v. a. centrata e normalizzata Y con

$$Y = \frac{x-m}{s} \qquad (2)$$

In tal caso,essendo l'integrando una funzione dispari,si ha:

$$E_y\{y\} = \int_{-\infty}^{\infty} \frac{y}{\sqrt{2p}} e^{-\frac{y^2}{2}} dy = 0 \qquad (3)$$

Ed inoltre per la varianza di y si ha:

$$s_y^2 = E_y\left\{(y-m_y)^2\right\} = \int_{-\infty}^{\infty} y^2 \frac{1}{\sqrt{2p}} e^{-\frac{y^2}{2}} dy = -\int_{-\infty}^{\infty} \frac{y}{\sqrt{2p}} de^{-\frac{y^2}{2}} = -\left[\frac{y}{\sqrt{2p}} e^{-\frac{y^2}{2}}\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2p}} dy = 1 \qquad (4)$$

Conseguentemente essendo :

$$X = s\ y + m \qquad (5)$$

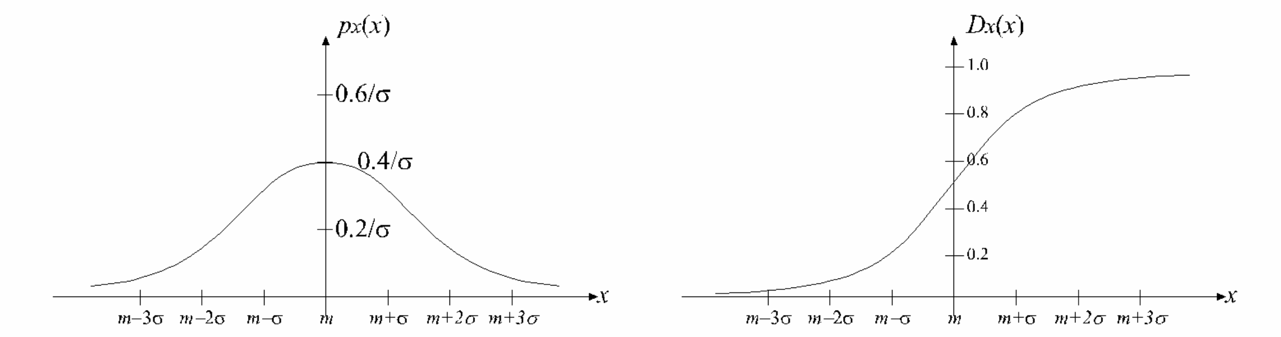
Si ha:

$$m_x = E_x\{x\} = E_y\{s \cdot y\} + m = m \qquad (6)$$

$$s^{\,2} = E_x\left\{(x - m_x)^2\right\} = E_y\{s \cdot y\} = s^{\,2} E_y\left\{y^2\right\} = s^{\,2} \qquad (7)$$

c.v.d.

La f.d.p. gaussiana unidimensionale e la relativa funzione di distribuzione sono riportate in fig.1a e 1b rispettivamente.



Come si vede da tale figura, la f.d.p. è simmetrica rispetto a m e unimodale (cioè con un solo massimo). Inoltre la probabilità che la v.a. cada nell' intervallo [m- **S** ,m+ **S**] è pari a 0.683.

Analogamente la probabilità che la v.a. x cada nell'intervallo [m-2 **S** ,m+2 **S**] è pari a 0.954 mentre,

la probabilità che essa cada nell'intervallo [m-3 **S** ,m+3 **S**] è pari a 0.977.Per tale ragione le specifiche degli errori di picco sono convertite in valori a 3 **S** nei casi pratici qualora venga adottato un modello gaussiano per tali errori.

2.1-FUNZIONE CARATTERISTICA

Per il calcolo della funzione caratteristica Px(?) si può procedere considerando ancora la v.a.

centrata e normalizzata y per la quale si ha: $y = \frac{m + x}{s}$

$$P_y(x) = \int_{-\infty}^{\infty} e^{-j2pxy} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2p}} dy = \int_{-\infty}^{\infty} \frac{e^{-\frac{y^2 - j4pxy - 4p^2x^2}{2}}}{\sqrt{2p}} e^{-2p^2x^2} dy = e^{-2p^2x^2} \int_{-\infty}^{\infty} \frac{e^{-\frac{(y - j2px)^2}{2}}}{\sqrt{2p}} dy = e^{-2p^2x^2} \qquad (8)$$

Pertanto poiché x = s y-m si ha :

$$P_x(x) = \Im\left\{\frac{1}{s\sqrt{2p}} e^{-\frac{(x-m)^2}{2s^2}}\right\} = \frac{1}{s}\Im\left\{\frac{e^{-\frac{(x-m)^2}{2s^2}}}{\sqrt{2p}}\right\} = \frac{s}{s} P_y(sx) e^{-j2pxm} = e^{-j2pxm} e^{-2p^2x^2s^2} \qquad (9)$$

2.2- FUNZIONE D'ERRORE

La funzione di distribuzione **Dx(x)** non può essere espressa per mezzo di funzioni elementari. Pertanto in passato sono stati sviluppati algoritmi per il calcolo numerico di tale funzione,nonché espressioni asintotiche.

In generale, nella letteratura, più che alla funzione di distribuzione si fa riferimento alla **funzione d' Errore.** La definizione di tale funzione non è universale,pertanto occorre ogni volta controllare la definizione adottata dal singolo autore. La definizione qui adottata è la seguente :

Funzione d' errore

$$erf\left(x\right)\overset{\Delta}{=}\frac{2}{\sqrt{p}}\int_{0}^{x} e^{-t^{2}}dt \qquad (10)$$

Funzione d' errore complementata

$$erfc\left(x\right)=1-erf\left(x\right)=\frac{2}{\sqrt{p}}\int_{x}^{\infty} e^{-t^{2}}dt \qquad (11)$$

In base a tali definizioni si ha che

$$D_{x}\left(x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2p}\,s}e^{-\frac{\left(t-m\right)^{2}}{2s^{2}}}dt=\frac{1}{2}\left[1+erf\left(\frac{x-m}{\sqrt{2}s}\right)\right]=1-\frac{1}{2}erfc\left(\frac{x-m}{\sqrt{2}s}\right) \qquad (12)$$

Una buona approssimazione della funzione d' errore complementata, per grandi valori dell' argomento è data da

$$erfc\left(x\right)\cong\frac{1}{\sqrt{p}\,x}e^{-x^{2}} \qquad (13)$$

Tale approssimazione asintotica può essere ottenuta integrando per parti la (11) ottenendo

$$erfc\left(x\right)=-\frac{2}{\sqrt{p}}\int_{x}^{\infty}\frac{1}{2t}de^{-t^{2}}=\frac{1}{\sqrt{p}}e^{-x^{2}}-\frac{1}{\sqrt{p}}\int_{x}^{\infty}\frac{1}{t^{2}}e^{-t^{2}}dt \qquad (14)$$

Poiché il secondo integrando è ovviamente positivo si ha

$$erfc\left(x\right)\leq\frac{1}{\sqrt{p}\,x}e^{-x^{2}} \qquad (15)$$

Ed inoltre il contributo trascurato tende a zero al tendere di x a $+\infty$

Migliori approssimazioni, ovviamente più complesse,sono riportate in [BURJESSON e SSUNDBERG,1979].

2.3-MOMENTI

I momenti centrati $m_x^{(11)}$ di ordine n di una v.a.gaussiana valgono :

$$m_x^{(11)} = \begin{cases} 0 & \text{n dispari} \\ 1 \cdot 3 \cdot 5 \cdot .... \cdot (n-1) s^{\,n} & \text{n pari} \end{cases} \tag{16}$$

Una dimostrazione della precedente relazione può essere ottenuta tramite l'impiego della funzione caratteristica della v.a. centrata e normalizzata y data dalla (2) osservando che :

$$E\{y^n\} = \frac{1}{(-j2p)^n} \frac{d^n P_y(x)}{dx^n}\Big|_{x=0} = \frac{1}{(-j2p)^n} \frac{d^n e^{-2p^2 x^2}}{dx^n}\Big|_{x=0} = \begin{cases} 0 \\ 1 \cdot 3 \cdot 5 \cdot .... \cdot (n-1) \end{cases} \tag{17}$$

da cui ricordando che

$$m_x^{(n)} = E_x\{(x - m_x)^n\} = E_y\{s^n y^n\} = s^n E_y\{y^n\} \tag{18}$$

segue immediatamente la (13) c.v.d.

3-DISTRIBUZIONE GAUSSIANA N-DIMENSIONALE

DEF-2 : Una variabile aleatoria $\underline{x}$ N-dimensionale si dice gaussiana, o a distribuzione normale, se la sua f.d.p. congiunta è del tipo

$$P_{\underline{x}}(\underline{x}) = \frac{1}{(2p)^{\frac{N}{2}}\sqrt{\det[k_x]}} e^{-\frac{1}{2}(\underline{x}-\underline{m}_x)^T k_x^{-1}(\underline{x}-\underline{m}_x)} \quad (19)$$

in cui Kx è una matrice (NxN) definita positiva.

La condizione che Kx sia definita positiva assicura l' esistenza dell'inversa K_x^{-1} e la convergenza dell' integrale della f.d.p. . Una definizione più generale che include il caso in cui Kx sia semidefinita positiva può essere ottenuta facendo ricorso alla funzione caratteristica e verrà introdotta nel seguito.

Il vettore $\underline{m}_x$ e la matrice Kx rappresentano rispettivamente il valore atteso e la matrice di covarianza della v.a. (come verrà dimostrato nel seguito).

Una notazione comune per indicare un v.a. gaussiana è la seguente :

$$\underline{x} \approx N(\underline{m}_x, \underline{k}_x) \quad (20)$$

Prima di dimostrare che $\underline{m}_x$ e Kx rappresentano il valore atteso e la matrice di covarianza della v. a. è utile dimostrare la seguente proprietà

PROPRIETA' 1- Una trasformazione lineare invertibile $\underline{Y}=A\underline{X}$ di un av.a. gaussiana è ancora gaussiana con

$$\underline{m_y} = A\underline{m_x} \quad (21.a)$$

$$K_y = AK_xA^T \quad (21.b)$$

PROVA- Per dimostrare tale proprietà si può applicare direttamente la regola del cambiamento di variabile e osservando che il determinante jacobiano è proprio pari a $\det\left[A^{-1}\right]$ si ha :

$$P_y(y)=P_x\left(A^{-1}y\right)\cdot\left|\det\left[A^{-1}\right]\right|=\frac{e^{-\frac{1}{2}\left(A^{-1}\underline{y}-\underline{m}_x\right)^T K_x^{-1}\left(A^{-1}\underline{y}-\underline{m}_x\right)}}{(2p)^{N/2}\sqrt{\det[K_x](\det[A])^2}}=\frac{e^{-\frac{1}{2}\left(y-A\underline{m}_x\right)^T\left(A^{-1}\right)^T K_x^{-1}A^{-1}\left(\underline{y}-A\underline{m}_x\right)}}{(2p)^{N/2}\sqrt{\det[K_x](\det[A])^2}} \qquad (22)$$

ovvero ponendo $m_y = A\underline{m}_x$, $K_y = AK_xA^T$ e ricordando che

$$K_y^{-1}=\left(AK_xA^T\right)^{-1}=\left(A^{-1}\right)^T K_x^{-1}A^{-1} \qquad (23.a)$$

$$\det\left[K_y\right]=\det\left[K_x\right]\cdot(\det[A])^2 \qquad (23.b)$$

si ha che :

$$P_y(y)=\frac{1}{(2p)^{N/2}\sqrt{\det\left[K_y\right]}}e^{-\frac{1}{2}\left(\underline{y}-\underline{m}_y\right)^T K_y^{-1}\left(\underline{y}-\underline{m}_y\right)} \qquad (24)$$

c.v.d.

Per il calcolo del valore atteso $\underline{m}_x$ e della matrice di covarianza K_x conviene fare riferimento inizialmente ad una variabile aleatoria centrata e normalizzata $\underline{y}_o$:

$$\underline{y}_o \sim N(\underline{O}, I) \qquad (25)$$

In tal caso la f.d.p. $P_{\underline{y}_o}(\underline{y}_o)$ diviene :

$$P_{\underline{y}_0}(\underline{y}_0) = \frac{1}{(2p)^{N/2}} e^{-\frac{1}{2}\underline{y}_0^T \underline{y}_0} = \prod_{i=1}^{N} \frac{1}{\sqrt{2p}} e^{-\frac{y_{oi}^2}{2}} \qquad (26)$$

Da cui si ha immediatamente che, essendo le v. a. marginali y_{oi} mutuamente statisticamente indipendenti con f.d.p. gaussiana a valore atteso nullo e varianza unitaria, il valore atteso di $\underline{y}_o$ è nullo e la sua matrice di covarianza è pari alla matrice identità I :

$$E_{\underline{y}_o}\{y_{o_i}\} = 0 \qquad (27.a)$$

$$E_{\underline{y}_o}\{y_{o_i} y_{o_j}\} = 0 \forall_{i,j} | i \neq j \qquad (27.b)$$

ovvero

$$\underline{m}_{y_o} = E_{y_o}\{\underline{y}_o\} = 0 \qquad (28.a)$$

$$K_{y_o} = E_{y_o}\{(\underline{y}_o - \underline{m}_{y_o})(\underline{y}_o - \underline{m}_{y_o})^T\} = I \qquad (28.b)$$

Ciò posto per derivare il valore atteso e la matrice di covarianza della v.a. $\underline{x}$ è utile introdurre la seguente proprietà nota come **Teorema di decomposizione di Cholesky :**

TEOREMA1- (Decomposizione di Cholesky) Data una matrice A definita non negativa, esiste ed è unica una matrice B triangolare in basso con elementi positivi o nulli sulla diagonale principale tale che :

$$A = B^T B \qquad (29)$$

La dimostrazione di tale teorema è riportata in appendice insieme all' algoritmo per il calcolo della matrice B.

In virtù di tale teorema data una v.a. $\underline{x}$:

$$\underline{x} \sim N(\underline{m}_x, K_x) \qquad (30)$$

da essa è sempre possibile,essendo K_x e quindi K_x^{-1} definita positiva , derivare tramite la trasformazione lineare :

$$\underline{y}_o = B(\underline{x} - \underline{m}_o) \qquad (31)$$

Con

$$K_x^{-1} = B^T B \qquad (32)$$

Una v.a. $\underline{y}_o \sim N(\underline{O}, I)$ infatti in base alla (21.b) si ha :

$$K_{y_0} = BK_x B^T = B\left(K_x^{-1}\right)^{-1} B^T = B\left(B^T B\right)^{-1} B^T = BB^{-1} B^{T^{-1}} B^T = I \qquad (33)$$

Poiché quindi

$$\underline{x} = B^{-1}\underline{y}_o + \underline{m}_x \qquad (34)$$

Ricordando le proprietà delle trasformazioni linari si ha:

$$E_{\underline{x}}\{\underline{x}\} = E_{\underline{y}_o}\left\{B^{-1}\underline{y}_o\right\} + \underline{m}_x = \underline{m}_x \qquad (35)$$

e

$$E_x\left\{(\underline{x} - \underline{m}_x)(\underline{x} - \underline{m}_x)^T\right\} = B^{-1} E_{\underline{y}_o}\left\{\underline{y}_o \underline{y}_o^T\right\}\left(B^{-1}\right)^T = B^{-1} I \left(B^{-1}\right)^T = \left(B^{-1}\right)\left(B^{-1}\right)^T = \left(B^T B^{-1}\right) = \left(K_x^{-1}\right) = K_x \qquad (36)$$

c.v.d.

Si noti che la trasformazione (29) è tale da trasformare un avariabile aleatoria gaussiana $\underline{x}$ con valore atteso $\underline{m}_x$ e matrice di covarianza K_x in una v. a. gaussiana con componenti centrate,normalizzate e incorrelate. Per tale motivo l' operatore B è noto nella letteratura con il termine di matrice (o filtro) sbiancante.

TEOREMA2-Qualora le variabili aleatorie risultino incorrelate la matrice di covarianza K_X diviene diagonale, ovvero

$$K_{\underline{x}} = \begin{bmatrix} s_{x_1}^2 & & \ldots & 0 \\ 0 & s_{x_2}^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & s_N^2 \end{bmatrix} \qquad (37)$$

conseguentemente anche la sua inversa K_x^{-1} è digonale :

$$K_x^{-1} = \begin{bmatrix} \frac{1}{s_{x_1}^2} & 0 & \ldots & 0 \\ 0 & \frac{1}{s_{x_2}^2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \frac{1}{s_{x_N}^2} \end{bmatrix} \qquad (38)$$

e la f.d.p.congiunta si riduce a

$$p_{\underline{x}}(\underline{x}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2p}\, s_{x_i}} e^{-\frac{(x_i - m_{x_i})^2}{2s_{x_i}^2}} \qquad (39)$$

per cui risulta dimostrato il seguente teorema.

TEOREMA2-Variabili aleatorie congiuntamente gaussiane incorrelate risultano anche statisticamente indipendenti .

Si noti che in base alla definizione di f.d.p. gaussiana N-dimensionale, il luogo dei punti dello spazio per cui la densità di probabilità risulta costante è costituito dai punti per i quali risulta costante l' esponente dell' equazione ( ) ovvero per i quali

$$\left(\underline{x}-\underline{m}_x\right)^T K_x^{-1}\left(\underline{x}-\underline{m}_x\right)=C^{'}\rangle 0 \qquad (40)$$

Tale equazione rappresenta quindi un iperellissoide con centro nel punto $\underline{m}_x$ i cui semiassi principali sono proporzionali alla radice quadrata degli autovalori della matrice K_x (che coincidono con gli inversi degli autovalori della matrice di covarianza K_x^{-1}. Inoltre le direzioni degli assi principali coincidono con quelle degli autovettori di K_x (ovvero di K_x^{-1}).Pertanto la matrice di covarianza determian la grandezza e l' orientamento di tali iperellissoidi che chiameremo "iperellissoidi di concentrazione". Se le variabili aleatorie marginali sono incorrelate K_x è diagonale e quindi gli assi principali degli iperellissoidi sono paralleli agli assi coordinati ed i semiassi sono proporzionali ai valori $s_i = \sqrt{K_{ii}}$.

A questo punto si può comprendere come considerando una trasformazione di coordinate che assuma come nuova base quella individuata dagli assi principali degli iperellissoidi di concentrazione,e che consiste quindi in una semplice rotazione, si ottenga una nuova v.a. ancora gaussiana ma v. a. marginali incorrelate e con varianze proporzionali agli autovalori di K. Tale proprietà può essere dimostrata in modo rigoroso facendo ricorso alle procedure di diagonalizzazione di forme quadratiche definite positive. Si noti che tale trasformazione non coincide con quella individuata dal filtro sbiancante. D' altronde la trasformazione lineare che diagonalizza una forma quadratica non è unica e per poterla individuare in modo univoco occorre porre ulteriori vincoli che nel caso in questione sono identificabili in :

a) filtro sbiancante : Ttrasformazione con matrice triangolare in basso
b) rotazione : trasformazione con matrice Q tale che

$$QQ^T = Q^T Q = I$$

3.1-**FUNZIONE CARATTERISTICA**

Anche per il calcolo della funzione caratteristica $P_{\underline{x}}(\xi)$ si può procedere considerando inizialmente il caso di una v.a. con componenti centrate, normalizzate e incorrelate ed applicare successivamente la decomposizione di Cholesky.

Data una v.a. $\underline{Y}_0$~N($\underline{O}$,I), poiché le singole componenti sono statisticamente indipendenti in base alla (8) si ha :

$$P_{\underline{y}_0}(\underline{x}) = E_{y_0}\left\{e^{-j2p\underline{x}^T\underline{y}_0}\right\} = E_{y_0}\left\{\prod_{i=1}^{N} e^{-j2pxy_{0_i}}\right\} = \prod_{i=1}^{N} e^{-2p^2x_i^2} = e^{-2p^2\sum_{i=1}^{N}x_i^2} = e^{-2p^2\underline{x}^T\underline{x}} \qquad (41)$$

pertanto,poiché $\underline{x} = B^{-1}\underline{y}_0 + m_x$ e $\det(B^{-1}) = (\det[K_x])^{\frac{1}{2}}$

si ha :

$$P_{\underline{x}}(\underline{x}) = \Im\left\{\frac{1}{(2p)^{N/2}\sqrt{\det[K_x]}} e^{-\frac{1}{2}(x-m_x)^T K_x^{-1}(\underline{x}-\underline{m}_x)}\right\} = \frac{1}{\sqrt{\det[K_x]}}\Im\left\{\frac{1}{(2p)^{N/2}} e^{-\frac{1}{2}(\underline{x}-\underline{m}_x)^T B^T B(\underline{x}-\underline{m}_x)}\right\} =$$

$$= e^{-j2p\,\underline{m}_x^T\underline{x}}\, e^{-2p^2\underline{x}^T(B^{-1})(B^{-1})^T\underline{x}} = e^{-j2p\,\underline{m}_x^T\underline{x}}\, e^{-2p^2\underline{x}^T K_x\underline{x}} \qquad (42)$$

In base atale relazione, in cuicompare direttamente la matrice di covarianza K_x e non la sua inversa, è possibile introdurre la seguente definizione alternativa di v.a. gaussiana che estende la def.1 al caso in cui K_x sia definita non negativa.

DEF.2- Una variabile aleatoria $\underline{x}$ N-dimensionale si dice gaussiana, o a distribuzione normale se la sua funzione caratteristica è del tipo

$$P_{\underline{x}}(\underline{x}) = e^{-j2p\underline{x}^T\underline{m}_x}\, e^{-2p^2\underline{x}^T K_x\underline{x}} \qquad (43)$$

con K_x matricce definita non negativa .

Tramite la funzione caratteristica , è possibile estendere facilmente la proprietà 1 al caso di trasformazione lineare $\underline{y}=A\underline{x}$ qualsiasi.

PROPRIETA' 1'- Una trasformazione lineare $\underline{y}=A\underline{x}$ di una v. a. gaussiana X~N ($\underline{m}_x$, K_x) è ancora gaussiana con

$$\underline{m}_y = A\underline{m}_x \tag{44.a}$$

$$K_y = AK_xA^T \tag{44.b}$$

Prova- Per definizione la funzione caratteristica $P_{\underline{y}}(\underline{\xi})$ di $\underline{y}$ è pari a :

$$P_{\underline{y}}(\underline{x}) = E_{\underline{y}}\left\{e^{-j2p\underline{x}^T\underline{y}}\right\} = E_{\underline{x}}\left\{e^{-j2p\underline{x}^TA\underline{x}}\right\} = E_x\left\{e^{-j2p\left(A^T\underline{x}\right)^T\underline{x}}\right\} = P_{\underline{x}}\left(A^T\underline{x}\right) = e^{-j2p\underline{m}_x^TA^T\underline{x}}e^{-2px^TAK_xA^Tx} \tag{45}$$

E quindi $\underline{y}$ è una v. a. gaussiana a valor atteso e matrice di covarianza data dalle (44.a) e (44.b) rispettivamente.

Utilizzando la precedente proprietà è facile dimostrare inoltre la proprietà seguente.

PROPRIETA' 2- ogni variabile marginale $\underline{x}_M$ M-dimensionale di una v. a. $\underline{x}$ gaussiana N-dimensionale è ancora gaussiana.

Prova: Per la dimostrazione di tale proprietà si può assumere, senza perdita di generalità che $\underline{x}_M$ sia costituito dalòle prime M componenti di $\underline{X}$ e che quindi $\underline{X}$ possa essere espresso come segue

$$\underline{x} = \begin{bmatrix} \underline{x}_M \\ \underline{x}_{N-M} \end{bmatrix} \tag{46}$$

in cui $\underline{X}_{N-M}$ rappresenta la v.a. marginale complementare a $\underline{X}_M$.

In tal caso è possibile esprimere il vettore dei valori attesi e la matrice di covarianza in forma ripartita come segue :

$$\underline{m}_x = \begin{bmatrix} \underline{m}_{x_M} \\ \underline{m}_{x_{N-M}} \end{bmatrix} \qquad (47.a)$$

$$K_x = \begin{bmatrix} K_{x_M} & K_{x_M x_{N-M}} \\ K_{x_{N-M} x_M} & K_{x_{N-M}} \end{bmatrix} \qquad (47.b)$$

Poiché $\underline{X}$ è gaussiana la v.a. $\underline{X}_M$

$$\underline{x}_M = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} \underline{x}_M \\ \underline{x}_{N-M} \end{bmatrix} \qquad (48)$$

è ancora gaussiana ,in virtù della proprietà 1' con :

$$E\{\underline{x}_M\} = \underline{m}_{x_M} \qquad (49.a)$$

$$E\left\{(\underline{x}_M - \underline{m}_x)(\underline{x}_M - \underline{m}_x)^T\right\} = IK_{x_M} I^T + \underline{O} = K_{x_M} \qquad (49.b)$$

c.v.d.

Commento. Si noti che tale proprietà definisce una condizione necessaria per le v.a. marginali ma non una condizione sufficiente. Per convincersi di ciò si noti che la funzione caratteristica della v.a. marginali $\underline{X}_i$ si ottiene dalla $P_{\underline{X}}(\underline{\xi})$ ponendo $\xi_{ij}=0$ per j≠i e che quindi aggiungendo all' esponente della funzione caratteristica data dalla ( ) un termine in $x_{i_i}^2 \quad x^2{}_{i_j}$ si ottiene una distribuzione che non è normale N-dimensionale ma possiede distribuzioni marginali gaussiane.

Un' ulteriore proprietà delle v.a. gaussiane che può essere derivata facendo ricorso alla funzione caratteristica e che riguarda il calcolo del momento congiunto del quarto ordine è quella espressa dal seguente teorema per la cui dimostrazione si rimanda all' appendice.

TEOREMA- Siano X_1, X_2, X_3, X_4 v.a. congiuntamente gaussiane con valori attesi m_{X_i} , i=1…4 e matrice di covarianza K_x con elementi $K_{xij}=\sigma_{x_ix_j}$, i,j =1…4. Allora :

$$m_x^{(1,1,1,1)} = E_x\left\{(x_1 - m_{x_1})(x_2 - m_{x_2})(x_3 - m_{x_3})(x_4 - m_{x_4})\right\} = s_{x_1x_2} s_{x_3x_4} + s_{x_1x_3} s_{x_2x_4} + s_{x_1x_4} s_{x_2x_3} \tag{50}$$

Si noti che in tale teorema le v.a. non devono essere necessariamente tutte distinte e quindi come caso particolare della ( ) si ottiene che sotto lev stesse ipotesi :

$$E_x\left\{(x_1 - m_{x_1})^2 (x_2 - m_{x_2})^2\right\} = s_{x_1}^2 s_{x_2}^2 + 2\left(s_{x_1x_2}^2\right)^2 \tag{51}$$

Un utile corollario del precedente teorema è quello che riguarda il momento centrale del terzo ordine.

<u>Corollario</u> – Il momento centrale congiunto di ordine 3 di 3 v. a. congiuntamente gaussiane è nullo.

$$E_x\left\{(x_1 - m_{x_1})(x_2 - m_{x_2})(x_3 - m_{x_3})\right\} = 0 \tag{52}$$

3.2- FUNZIONI DI DENSITA' DI PROBABILITA' CONDIZIONATA

Nel paragrafo precedente è stato dimostrato che date due variabili aleatorie $\underline{X}$ e $\underline{Y}$ congiuntamente gaussiane ripettivamente a m e n dimensioni, le loro distribuzioni marginali risultano gaussiane.Una ulteriore proprietà, estremamente utile nelle applicazioni, è che la v.a. $\underline{X}$ condizionata alla v.a. $\underline{Y}$ risulta ancora gaussiana. Tale proprietà può essere formalizzata come segue.

PROPRIETA'- Data una v.a. gaussiana $\underline{Z}$:

$$\underline{Z} = \begin{bmatrix} \underline{X} \\ \underline{Y} \end{bmatrix} \tag{50}$$

Con valore atteso $\underline{m}_Z$ e matrice di covarianza $\underline{K}_Z$:

$$\underline{m}_z = \begin{bmatrix} \underline{m}_x \\ \underline{m}_y \end{bmatrix} \tag{51.a}$$

$$K_z = \begin{bmatrix} K_x & K_{xy} \\ K_{yx} & K_y \end{bmatrix} \tag{51.b}$$

La f.d.p. di $\underline{X}$ condizionata ad $\underline{Y}$ $P_{\underline{X}/\underline{Y}}(\underline{x}/\underline{y})$ risulta guassiana con valore atteso $\underline{m}_{x/y}$ e matrice di covarianza $K_{x/y}$ dati da :

$$\underline{m}_{x/y} = \underline{m}_x + K_{xy} K_y^{-1} \left(\underline{y} - \underline{m}_y \right) \tag{52.a}$$

$$K_{x/y} = K_{xx} - K_{xy} K_y^{-1} K_{yx} \tag{52.b}$$

<u>Prova</u> : Per dimostrare tale proprietà si osservi che in base alla definizione la f.d.p. $P_{\underline{X}/\underline{Y}}(\underline{x}/\underline{y})$ può essere espressa come segue :

$$P_{\underline{x}/\underline{y}}\left(\underline{x}/\underline{y}\right)=\frac{P_{\underline{x},\underline{y}}\left(\underline{x},\underline{y}\right)}{P_{\underline{y}}\left(\underline{y}\right)}==\frac{1}{\dfrac{(2p)^{n+m/2}\sqrt{\det\left[K_z\right]}}{(2p)^{n/2}\sqrt{\det\left[K_y\right]}}}e^{-\frac{1}{2}\left[(\underline{z}-\underline{m}_z)^T K_z^{-1}(\underline{z}-\underline{m}_z)-(\underline{y}-\underline{m}_y)K_y^{-1}(\underline{y}-\underline{m}_y)\right]} \tag{53}$$

Si osservi che tale f.d.p. può anche essere espressa nella forma :

$$P_{\underline{x}/\underline{y}}\left(\underline{x}/\underline{y}\right)=\frac{1}{(2p)^{m/2}\sqrt{\det\left[K_{x/y}\right]}}e^{-\frac{1}{2}(\underline{x}-\underline{m}_{x/y})^T K_{x/y}^{-1}(\underline{x}-\underline{m}_{x/y})} \tag{54}$$

In cui $\underline{m}_{x/y}$ e $K_{x/y}$ sono rispettivamente il valore atteso e la matrice di covarianza condizionate.

Eguagliando le due espressioni e osservando che la matrice K_z^{-1} può essere espressa come segue :

$$K_z^{-1}=\begin{bmatrix}\left(K_x-K_{xy}K_y^{-1}K_{yx}\right)^{-1} & -K_x^{-1}K_{xy}\left(K_y-K_{yx}K_{yx}K_x^{-1}K_{xy}\right)^{-1}\\ -K_y^{-1}K_{yx}\left(K_x-K_{xy}K^{-1}K_{yx}\right)^{-1} & \left(K_y-K_{yx}K_x^{-1}K_{xy}\right)^{-1}\end{bmatrix} \tag{55}$$

Si ha immediatamente che :

$$K^{-1}{}_{x/y}=\left(K_x-K_{xy}K_y^{-1}K_{yx}\right)^{-1} \tag{56.a}$$

$$K_{x/y}^{-1}\underline{m}_{x/y}=\left(K_xK_{xy}K^{-1}K_{yx}\right)^{-1}\underline{m}_x+K_x^{-1}K_{xy}\left(K_y-K_{yx}K_x^{-1}K_{xy}\right)^{-1}\left(\underline{y}-\underline{m}_y\right) \tag{56.b}$$

Osservando che per il lemma di inversione matriciale si ha :

$$K_x^{-1}K_{xy}\left(K_y-K_{yx}K_x^{-1}K_{xy}\right)^{-1}=\left(K_x-K_{xy}K_y^{-1}K_{yx}\right)^{-1}K_{xy}K_y^{-1} \tag{57}$$

Infine si ha :

$$K_{x/y} = K_x - K_{xy} K_y^{-1} K_{yx} \tag{58.a}$$

$$\underline{m}_{x/y} = \underline{m}_x + K_{xy} K_y^{-1} \left(y - \underline{m}_y \right) \tag{58.b}$$

c.v.d.

- DISTRIBUZIONE GAUSSIANA BIDIMENSIONALE

Per il suo utilizzo frequente è utile esplicitare le relazioni trovate in precedenza nel caso bidimensionale. Indicata quindi con Z≡(x,y) una v.a. bidimensionale con v.a. marginali X e Y sia

$m_x \stackrel{\Delta}{=} E_x\{x\}$: valore atteso di x

$m_y \stackrel{\Delta}{=} E_y\{y\}$: valore atteso di y

$s_x^2 \stackrel{\Delta}{=} E_x\{(x-m_x)^2\}$: varianza di x

$s_y^2 \stackrel{\Delta}{=} E_y\{(y-m_y)^2\}$: variazione di y

$s_{xy} \stackrel{\Delta}{=} E_{x,y}\{(x-m_x)(y-m_y)\}$: covarianza di x e y

$r \stackrel{\Delta}{=} \frac{s_{xy}}{s_x s_y}$: coefficiente di correlazione tra x e y

da cui segue che :

$$m_z = \begin{bmatrix} m_x \\ m_y \end{bmatrix} \tag{59}$$

$$K_z = \begin{bmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{bmatrix} = \begin{bmatrix} s_x^2 & r s s_y \\ r s_x s_y & s_y^2 \end{bmatrix} \tag{60}$$

Pertanto per il determinante della matrice di covarianza si ha :

$$\det[K_z] = s_x^2 s_y^2 (1 - r^2) \tag{61}$$

E per la sua in versa risulta :

$$K_x^{-1} = \frac{1}{1-r^2}\begin{bmatrix} \dfrac{1}{s_x^2} & \dfrac{r}{-s_x s_y} \\ \dfrac{-r}{s s_{yx}} & \dfrac{1}{s_y^2} \end{bmatrix} \qquad (62)$$

Per cui la f.d.p. bidimensionale risulta del tipo :

$$P_{x,y}(x,y) = \frac{e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{s_x^2} - \frac{2r(x-m_x)(y-m_y)}{s_x s_y} + \frac{(y-m_y)^2}{s_y^2}\right]}}{2p s_x s_y \sqrt{1-r^2}} \qquad (63)$$

Mentre per la funzione caratteristica si ha :

$$P_{x,y}(x_1,x_2) = e^{-j2p(m_x x_1 + m_y x_2)} e^{-2p^2(s_x^2 x_1^2 + 2 r s_x s_y x_1 x_2 + s_y^2 x_2^2)} \qquad (64)$$

Dall’ esame della f.d.p. congiunta si deduce che il luogo dei punti del piano per cui la f.d.p è costante, ad esmpio pari ad un valore c, è costituito dai punti del piano per i quali risulta costante l’esponente dell’ equazione (63), ovvero dai punti (x,y) che soddisfano l’ equazione :

$$\frac{(x-m_x)^2}{s_x^2} - \frac{2r(x-m_x)(y-m_y)}{s_x s_y} + \frac{(y-m_y)^2}{s_y^2} = C' \qquad (65)$$

Che rappresenta un ellisse con centro nel punto (m_x, m_y) e assi principali ruotati di un angolo α

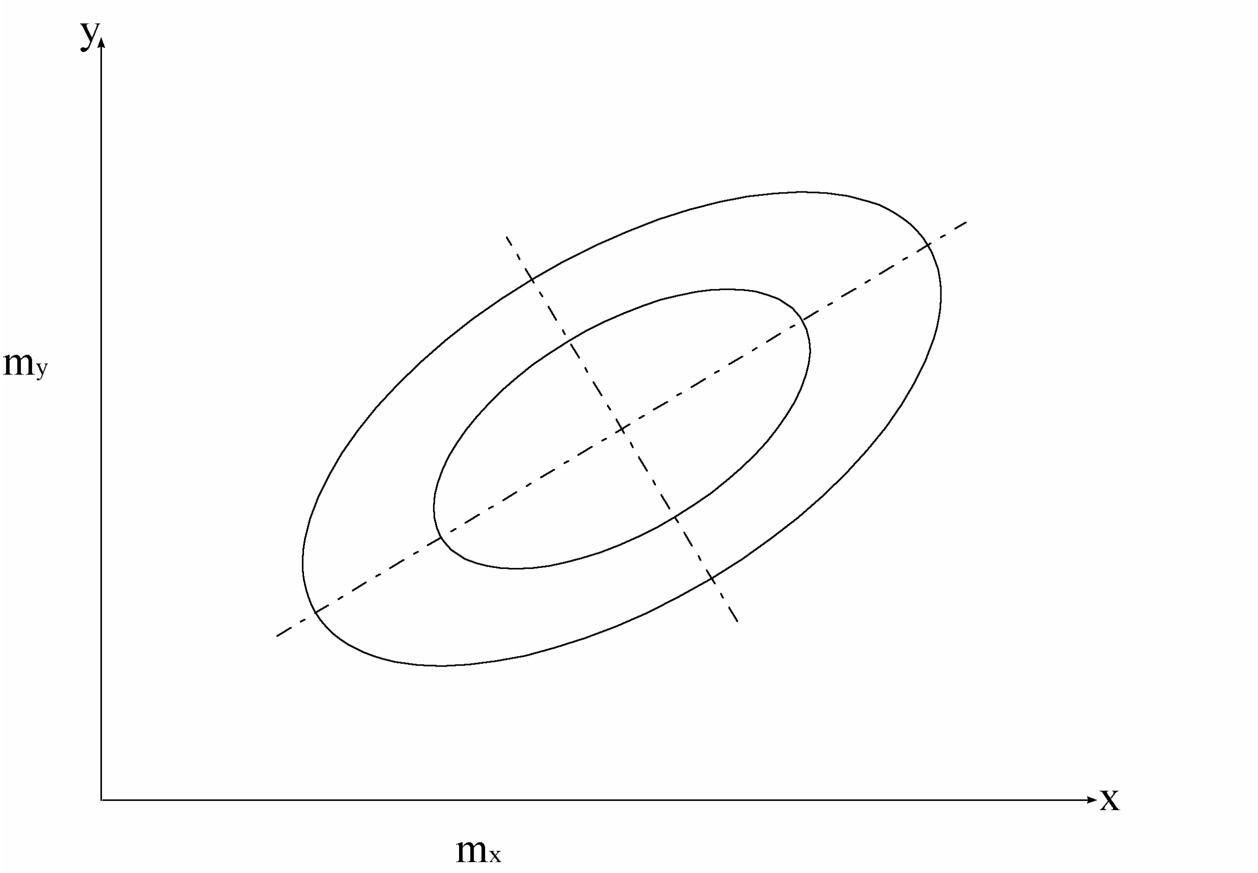


_fig.ellissi di concentrazione

Come riportato in fig. tali ellissi prendono anche il nome di ” ellissoidi di concentrazione”. E’ possibile dimostrare inoltre che le lunghezze dei semiassi principali sono inversamente proporzionali alla radice quadrata degli autovalori della matrice di covarianza K_z^{-1}. Nel caso in cui le due v.a. x ey abbiano la stessa varianza $s_x^2 = s_y^2 = s^2$ l’ equazione si riduce a

$$(x-m_x)^2 - 2r(x-m_x)(y-m_y) + (y-m_y)^2 = C' \qquad (C') 0 \qquad (66)$$

Pertanto salvo il caso di ρ=0 per cui le ellissi si riducono a cerchi con centro nel punto (m_x, m_y), gli assi principali coincidono con le parallele alle due bisettrici del 1° e 3° e del 2° e 4° quadrante (vedi fig.)

In tal caso è facile verificare che considerata la trasformazione

$$x_1 = \frac{1}{\sqrt{2}}(x - y) \qquad (67.a)$$

$$y_1 = \frac{1}{\sqrt{2}}(x + y) \qquad (67.b)$$

A cui corrisponde la matrice di trasformazione

$$A = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad (68)$$

Si ha in effetti che la matrice di covarianza K_{z_1} di $z_1 \equiv (x_1, y_1)$ diviene diagonale

$$K_{z_1} = AK_zA^T = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} s^2 & rs^2 \\ rs^2 & s^2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} (1-r)2s^2 & 0 \\ 0 & (1+r)2s^2 \end{bmatrix} \qquad (69)$$

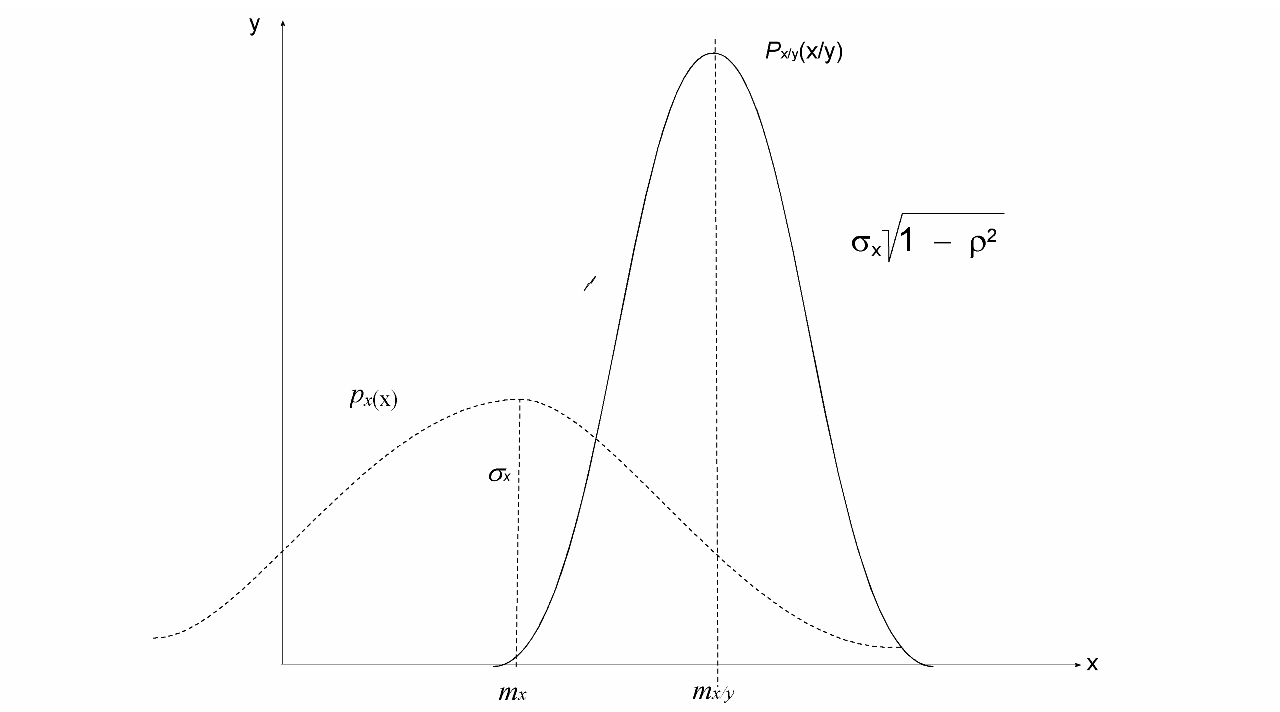
E quindi le lunghezze dei semiassi principali sono proporzionali a $(1-r)s$ (2^o e 4^0 quadrante) e $(1+r)s$ (1^0 e 3^0 quadrante).

Per le distribuzioni condizionate si ha che :

$$\begin{cases} m_{x/y} = m_x + r\dfrac{s_x}{s_y} & (70.a) \\ s^2_{y/x} = s^2_x\left(1 - r^2\right) & (70.b) \end{cases}$$

$$
\begin{cases} m_{y/x} = m_y + r\,\dfrac{s_y}{s_x}\left(x - m_x\right) & (70'.a) \\ s^2_{y/x} = s^2_y\left(1 - r^2\right) & (70'.b) \end{cases}
$$

Tale risultato può essere interpretato come segue. Se la f.d.p. della v.a. x è del tipo riportato in Figura



Allora la conoscenza che la v.a. y ha assunto la determinazione y riduce l' incertezza sulla v.a. x in Quanto la f.d.p. condizionata di x rispetto a y è ancora gaussaina ma il suo valore atteso m_x si modifica di una quantità proporzionale allo scostamento $(y-m_y)$

tra l'attuale determinazione della v.a. y ed il suo valore atteso, secondo un fattore pari a $r\,\dfrac{s_x}{s_y}$

mentre la sua varianza si riduce a $\left(1-r^2\right)s_x^2$. Si noti che tale riduzione è tanto più forte quanto più |ρ| tende a 1. Infatti $r = \pm 1$ implica una relazione lineare tra le variabili e quindi nota y la determinazione della v.a. x può essere calcolata immediatamente.

Ciò si riflette nel fatto che per |ρ| =1 la varianza condizionata $s_{x/y}$ vale 0 e quindi la f.d.p. tende ad un impulso matematico centrato in $m_{x/y}$.

Tale riduzione invece non si manifesta nel caso ρ =0, ed inafatti in tal caso le due v.a. aleatorie,essendo gaussiane, oltre che incorrelate risultano anche statisticamente indipendenti e quindi l' osservazione di una delle due non porta nessuna inforamazione sul comportamento dell' altra.

Si noti infine che la curva di regressione di un adelle due v.a. rispetto all' altra è una retta la cui equazione è data dalla (70.a)