Jointly Gaussian Random Variables/ Gaussian Random Vectors

We saw the univariate normal density $X \sim N(\mu, \sigma^2)$:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

If $X_1, \dots, X_n$ are independent $N(\mu_i, \sigma_i^2)$ random variables, their joint pdf is the product of their individual densities:

$$f(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right]$$

$=$

where $\underline{x} = [x_1, x_2, \ldots, x_n]$

The exponent ~~of~~ in the density is a quadratic form. We try to rewrite it using matrix-vector notations.

Assume:

$$C = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_n^2 \end{bmatrix}$$

$$= \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$$

and $\underline{\mu} = [\mu_1, \mu_2, \ldots, \mu_n]^T$

Then, the exponent

$$\sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{\sigma_i^2} =$$

~~So, these~~ Observe that

$$\det(C) = \sigma_1^2 \sigma_2^2 \cdots \sigma_n^2$$

Therefore:

$$f(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(C)}} \exp\left(-\frac{1}{2}(\underline{x}-\underline{\mu})^T C^{-1} (\underline{x}-\underline{\mu})\right)$$

Since $X_1, \ldots, X_n$ are independent, they are uncorrelated, so

$$\mathrm{Cov}(X) = \begin{bmatrix} \mathrm{Var}(X_1) & & 0 \\ & \ddots & \\ 0 & & \mathrm{Var}(X_n) \end{bmatrix} = C$$

Linear Transformation of independent Gaussians

Assume $\underline{X} = [X_1 \cdots X_n]^T$, where X_i's are independent $N(\mu_i, \sigma_i^2)$.

We wish to derive the joint pdf of $\underline{Y} = A\underline{X} + \underline{b}$, where $A \in \mathbb{R}^{n \times n}$ is an invertible matrix and $\underline{b} \in \mathbb{R}^n$. We have:

$$f_{\underline{Y}}(\underline{y}) = \frac{f_{\underline{X}}(A^{-1}(\underline{y} - \underline{b}))}{|\det(A)|}$$

Therefore, given

$$f_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C}} \exp\left(-\frac{(\underline{x}-\underline{\mu})^T C^{-1} (\underline{x}-\underline{\mu})}{2}\right)$$

$$f_{\underline{Y}}(\underline{y}) =$$

Therefore, by Choosing

$\underline{\mu}' = A\underline{\mu} + \underline{b}$, and $C' = ACA^T$

and noting that

$\det(C') = \det(ACA^T) = \det(A)\det(C)$

$\det(A^T) = |\det(A)|^2 \det(C)$

$$f_{\underline{Y}}(\underline{y}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(C')}} \exp\left(-\frac{(\underline{y}-\underline{\mu}')^T C'^{-1} (\underline{y}-\underline{\mu}')}{2}\right)$$

C' can be non-diagonal. It seems that the joint pdf of $\underline{Y}$ has the same form as the joint pdf of $\underline{X}$.

We call any vector $\underline{Y}$ that has the above joint pdf, a ~~jointly~~ Gaussian random vector, or alternatively $Y_1, \ldots, Y_n$ are

called jointly normal/Gaussian, with mean vector $\underline{\mu}$, and covariance matrix, C.
What happens if A is singular?

We previously saw that when A is singular, $\underline{Y} = A\underline{X} + \underline{b}$ is no longer jointly continuous.

However, each of Y_i's is still a Gaussian random variable, because we can show, using convolution integrals, that any linear combination of independent $\mathcal{N}(\mu_i, \sigma_i^2)$'s is Gaussian.

Exercise: Show that if $X_1, \dots, X_n$ are $\mathcal{N}(\mu_i, \sigma_i^2)$ and independent, then

$$\sum_{i=1}^{n} c_i X_i \text{ is } \mathcal{N}\left(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2\right).$$

(Use convolution integrals. First show for $n=2$ and then use induction)

However because A is singular, a conventional joint pdf doesn't exist for $\underline{Y}$. This can also be seen from the fact that

$$Cov(\underline{Y}) = \mathbb{E}\left[\left(\underline{Y} - \underline{A}\underline{\mu} - \underline{b}\right)\left(\underline{Y} - A\underline{\mu} - \underline{b}\right)^T\right]$$

$$= \mathbb{E}\left[A(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T A^T\right]$$

$$= A\,\mathbb{E}\left[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^T\right]A^T$$

$= ACA^T$ is singular.

This leads us to use a more general definition of joint Gaussianity that does not rely on joint pdfs.

Def: $\underline{X} = [X_1, \ldots, X_n]^T$ is said to be jointly Gaussian if every linear combination of X_i's

$$\sum_{i=1}^{n} a_i X_i$$

is a Gaussian random variable.

If $\mathbb{E}[\underline{X}] = \underline{\mu}$ and $\mathrm{Cov}(\underline{X}) = C$, we write $\underline{X} \sim N(\underline{\mu}, \underline{C})$

Exercise: Show that if

$\underline{X} \sim \mathcal{N}(\underline{\mu}, C)$ and $A \in \mathbb{R}^{p \times n}$

, $\underline{b} \in \mathbb{R}^p$, then if $\underline{Y} = A\underline{X} + \underline{b}$,

$\underline{Y} \sim \mathcal{N}(A\underline{\mu} + \underline{b}, ACA^T)$.

Note: For the definition to make

sense, when $\sigma_i = 0 \ \forall i$, we

assume the degenerate r.v.

$X = 0$ is also Gaussian.

For Gaussian Random Vectors

Uncorrelated implies Independent

If $\underline{X} = [X_1, \dots, X_n]$ is $\mathcal{N}(\underline{\mu}, C)$,

and X_i and X_j are uncorrelated

$C = \mathrm{Cov}(\underline{X}) = \mathrm{diag}(\sigma_1^2, \dots, \sigma_n^2)$.

It can be seen that

$f_{\underline{X}}(\underline{x})$ is decomposed as

$$f_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left(-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}\right),$$

i.e. X_i's are independent

Moment Generating Function and Characteristic Function of Gaussian Vectors

(Joint MGFs and CFs)

Assume that $\underline{X} \sim N(\underline{\mu}_X, C_X)$

$$M_{\underline{X}}(\underline{s}) = \mathbb{E}[e^{s^T \underline{X}}] \qquad \underline{s} \in \mathbb{R}^n$$

On the other hand, we know that $Y = s^T \underline{X} = s_1 X_1 + \cdots + s_n X_n$ is a linear combination of jointly Gaussian r.v.'s, so it is Gaussian, i.e. $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$

Therefore $M_Y(t) =$

Also: $M_{\underline{X}}(s) = \mathbb{E}[e^{\underline{s}^T \underline{X}}]$

$= \mathbb{E}[e^{Y}] = \mathbb{E}[e^{Yt}]\Big|_{t=1}$

$= M_Y(1)$

$\Rightarrow M_{\underline{X}}(s) = e^{\mu_Y t + \frac{\sigma_Y^2 t^2}{2}}\Big|_{t=1}$

$= e^{\mu_Y + \frac{\sigma_Y^2}{2}}$

$\mu_Y = \mathbb{E}[s^T \underline{X}] =$

$$\sigma_Y^2 = \mathbb{E}[(Y - \mu_Y)(Y - \mu_Y)]$$

since Y is a scalar,

$$\sigma_Y^2 = \mathbb{E}[(Y - \mu_Y)(Y - \mu_Y)^T]$$

$=$

Therefore

$$M_{\underline{X}}(s) = e^{\mu_Y + \frac{\sigma_Y^2}{2}} =$$

The CF of $\underline{X}$ is

$$\varphi_{\underline{X}}(\underline{\omega}) = M_{\underline{X}}(\underline{s})\Big|_{\underline{s}=j\underline{\omega}}$$

$$= e^{j\underline{\omega}^T \underline{\mu}_X + \frac{1}{2}(j\underline{\omega})^T C_X (j\underline{\omega})}$$

$$= e^{j\underline{\omega}^T \underline{\mu}_X - \frac{1}{2}\omega^T C_X \omega}$$

Note: The joint pdf only exists if C^{-1} exists. If C is not invertible, the joint pdf doesn't exist in its conventional sense, and one has to use a generalized joint pdf since $\underline{X}$ is no-longer jointly continuous.

Therefore, sometimes a jointly normal random vector is

defined as a random vector having the ~~char~~ joint CF

$$M_{\underline{X}}(\underline{\omega}) = e^{j\omega^T \underline{\mu}_X - \frac{1}{2}\omega^T C \omega}$$

Conditional pdfs

Assume that $\begin{bmatrix} \underline{X} \\ \hline \underline{Y} \end{bmatrix}$ is a random vector and is Gaussian. Then one can show

$$\underline{X} \mid \underline{Y} = \underline{y} \sim \mathcal{N}\left(\mathbb{E}[\underline{X} \mid \underline{Y} = \underline{y}], C_{\underline{X}\mid\underline{Y}}\right)$$

where

$$C_{X\mid Y} = C_X - S C_{YX}$$

in which S is the solution to the matrix equation

$$S C_Y = C_{XY}$$

and if C_Y is invertible, $S = C_{XY} C_Y^{-1}$

so

$$C_{X|Y} = C_X - C_{XY} C_Y^{-1} C_{YX}$$

and

$$\mathbb{E}[\underline{X} \mid \underline{Y} = \underline{y}] = S(\underline{y} - \underline{m}_Y) + \underline{m}_X$$

and if C_Y is invertible

$$\mathbb{E}[\underline{X} \mid \underline{Y} = \underline{y}] = C_{XY} C_Y^{-1} (\underline{y} - \underline{m}_Y) + \underline{m}_X$$

The Conditional pdf then becomes

$$f_{\underline{X}|\underline{Y}}(x|y) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C_{X|\underline{Y}}}} \times \exp\left[-\frac{1}{2}(\underline{x}-g(\underline{y}))^T C_{X|\underline{Y}}^{-1}(\underline{x}-g(\underline{y}))\right]$$

provided that $C_{X|\underline{Y}}$ is invertible.