1 The univariate gaussian distribution

1.1 The density function

A continuous random variable is said to have the gaussian distribution if its density function is given by:

$$f_X(x) = K \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

The constant K is obtained by verifying that:

$$\int_{\Re} K \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x = 1 \tag{1}$$

Recalling that $\mathrm{P}(X \in [a,b]) = \int_a^b f_X(x)\mathrm{d}x$, equation 1 translates 'it is sure that X takes a value between $-\infty$ and $+\infty$'.

The method to evaluate K is given in appendix, it is suggested to attempt to solve it as an exercise. The final form of the univariate gaussian is:

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

1.2 Features

The expectation (*ie* the mean) of a gaussian as given by equation 2 is μ:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\Re} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x = \mu$$

It is straightforward to obtain this result with an appropriate change of variable.

The variance of a gaussian is σ^2:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\Re} (x-\mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x = \sigma^2$$

This result can be obtained with an integration per parts (let $u = (x-\mu)$ and v' the rest).

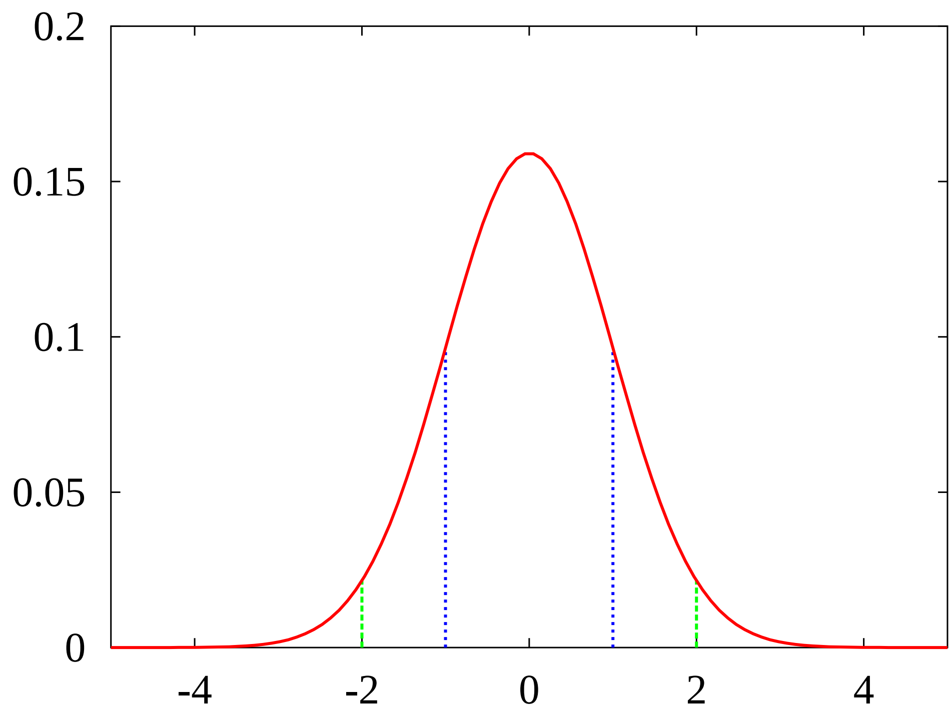

Figure 1: The standard ($\mu = 0$,$\sigma = 1$) gaussian, the probability to find X between $\mu - \sigma$ and $\mu + \sigma$ is about 0.67. It is of 0.95 between $\mu - 2\sigma$ and $\mu + 2\sigma$, and 0.99 between $\mu - 3\sigma$ and $\mu + 3\sigma$.

2 Multivariate standard gaussian

2.1 Bivariate standard gaussian

f is the standard bivariate gaussian (or normal) density function of two random variables X and Y if it is given by:

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right) \tag{3}$$

where ρ verifies $|\rho| < 1$. This can also be written:

$$f_{XY}(x,y) = \frac{\sqrt{|M|}}{2\pi} \exp\left(-\frac{1}{2}\mathbf{x}M\mathbf{x}^T\right)$$

where $\mathbf{x}$ refers to the vector (x, y) and M is the matrix:

$$M = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

You can easily verify that the determinant of M, denoted $|M|$ is $1/(1-\rho^2)$. The calculation of the normalisation constant is done later.

ρ as defined in equation 3 is the covariance of the variables X and Y. Recall $\text{cov}(X,Y) = E(XY) - E(X)E(Y)$ is a measure of how strongly the variables depend on each other. For example, weight and height of individuals in a population have a strong covariance.

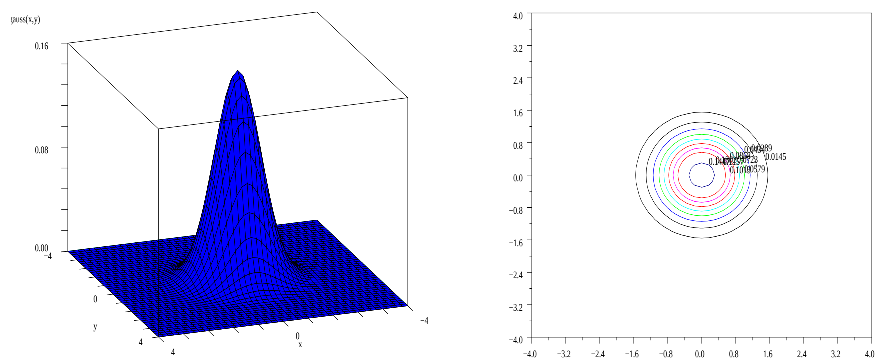

Figure 2: Standard bivariate gaussian, with $\rho = 0$, and corresponding contour plot.

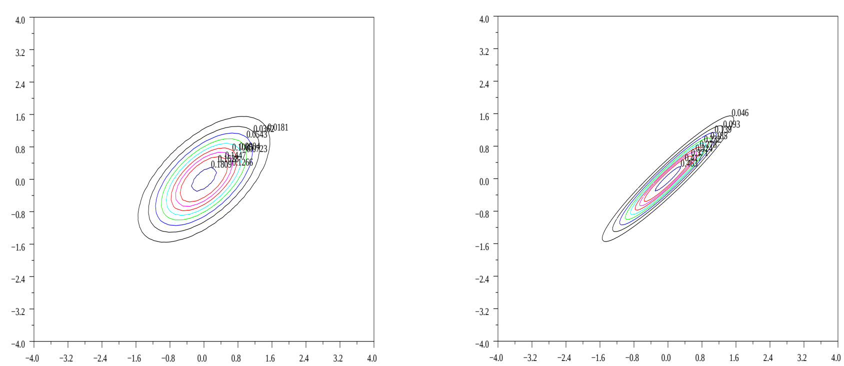

Figure 3: Contour plot for the standard bivariate with $\rho = 0.6$ and $\rho = 0.95$. As the covariance increases toward 1, the contour elongates along $x = y$, illustrating the fact that chosing a value for X or Y reduces the interval of 'possible' values for the other variable.

2.2 General case

2.2.1 The normalisation factor

In the general case, for n random variables, $\mathbf{x}$ denotes the vector $(x_1, x_2, \ldots, x_n)$. The notation is identical to the one seen previously, that is:

$$f(\mathbf{x}) = \sqrt{\frac{|M|}{(2\pi)^n}} \exp\left(-\frac{1}{2}\mathbf{x}M\mathbf{x}^T\right)$$

The proof for the value of the normalisation factor is now given: because $Q(\mathbf{x}) = \mathbf{x}M\mathbf{x}^T$ must be positive for all values of $\mathbf{x}$, Q is said to be definite positive. Under this condition, a theorem in matrix theory states that M can be diagonalised. What this means is that there exists a base in which M can be written:

$$M' = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Let us write $\mathbf{y}$ the vector $\mathbf{x}$ in the base in which M is diagonal (where it is denoted M'), and rewrite $Q(\mathbf{x})$:

$$Q(\mathbf{x}) = \mathbf{y}M'\mathbf{y}^T$$

the normalisation is now straightforward since the random variables are independent:

$$\begin{aligned}
1 &= K\int_{\Re^n} \exp\left(-\frac{1}{2}\mathbf{x}M\mathbf{x}^T\right) \mathrm{d}\mathbf{x} \\
&= K\int_{\Re^n} \exp\left(-\frac{1}{2}\mathbf{y}M'\mathbf{y}^T\right) \mathrm{d}\mathbf{y} \\
&= K\int_{\Re^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\lambda_i y_i^2\right) \mathrm{d}\mathbf{y} \qquad (4) \\
&= K\prod_{i=1}^{n}\int_{\Re} \exp\left(-\frac{1}{2}\lambda_i y_i^2\right) \mathrm{d}y_i \qquad (5) \\
& \qquad (6)
\end{aligned}$$

This imposes on K:

$$K = \frac{\sqrt{\lambda_1\lambda_2\ldots\lambda_n}}{\sqrt{(2\pi)^n}} = \frac{\sqrt{|M'|}}{\sqrt{(2\pi)^n}} = \frac{\sqrt{|M|}}{\sqrt{(2\pi)^n}} \qquad (7)$$

where $|M|$ denotes the *determinant* of M, which is independent of the basis in which M is written. As an exercise, diagonalise the standard bivariate gaussian given by 3 (see appendix for correction).

2.2.2 Usual form of the general multivariate gaussian

Most often, you will find the multivariate gaussian written:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |V|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\overline{\mathbf{x}})V^{-1}(\mathbf{x}-\overline{\mathbf{x}})^T\right) \tag{8}$$

where $\overline{\mathbf{x}}$ is the vector made up of the average (expectation) of each component of $\mathbf{x}$. In this form, V is the *covariance matrix*, made up of the v_{ij}:

$$v_{ij} = \mathrm{cov}(x_i, x_j) = \mathrm{cov}(x_j, x_i) = v_{ji}$$

Appendix

Evaluating the normalising constant for the univariate gaussian

HINT: The solution involves, first, a change of variable, and second, the evaluation of the square of the integral rather than the integral itself, this last step requires changing to polar coordinate.

SOLUTION: A first step consists in changing the variable to $u = (x - \mu)/\sigma$ so that:

$$K\sigma \int_{\Re} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1$$

To calculate $I = \int \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u$, a trick consists in writing:

$$I^2 = \int \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u \int \exp\left(-\frac{v^2}{2}\right) \mathrm{d}v = \iint \exp\left(-\frac{u^2+v^2}{2}\right) \mathrm{d}u\mathrm{d}v$$

Changing to polar coordinates ($r = x\cos\theta$, $y = r\sin\theta$), the element $\mathrm{d}u\mathrm{d}v$ becomes $r\mathrm{d}r\mathrm{d}\theta$ and $u^2 + v^2 = r^2$, giving:

$$I^2 = \int_0^{2\pi} \int_0^{+\infty} r\exp\left(-\frac{r^2}{2}\right) \mathrm{d}r\mathrm{d}\theta = 2\pi\left[-\exp(-r^2/2)\right]_0^{+\infty} = 2\pi$$

K must therefore be $1/\sqrt{(2\pi\sigma^2)}$ in order to verify (1)

Diagonalising the standard bivariate gaussian

The standard bivariate gaussian can be written:

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\frac{x^2+y^2-2\rho xy}{1-\rho^2}\right) = \frac{\sqrt{|M|}}{2\pi} \exp\left(-\frac{1}{2}\,\mathbf{x}M\mathbf{x}^T\right)$$

with

$$M = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

It is out of the scope of this lecture to explain diagonalisation. You may have a look at http://www.numbertheory.org/book/ (in general, this is basic matrix theory so any textbook should do). If you are not familiar with matrix theory, you might still convince yourself that, with:

$$x' = \frac{1}{\sqrt{2}}(x+y) \quad (9)$$

$$y' = \frac{1}{\sqrt{2}}(x-y) \quad (10)$$

and

$$M' = \frac{1}{1-\rho^2}\begin{bmatrix} 1-\rho & 0 \\ 0 & 1+\rho \end{bmatrix}$$

we have

$$(x,y)M(x,y)^T = (x',y')M'(x',y')^T = \frac{1}{2}\frac{x^2+y^2-2\rho xy}{1-\rho^2}$$