1 The Gaussian (Normal) Distribution

1.1 Introduction

The material presented in this chapter requires some elementary knowledge of probability and statistics. The symbol $f(\underline{x})$ denotes probability distribution assigned to an uncertain vector $\underline{x}$.

1.2 Standard Normal and General Normal Distribution

The probability density function (PDF) $f_Z(z) \equiv f(z)$ of a standard Gaussian variable or standard normal variable Z, denoted also by $\phi(z)$, is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) \tag{1}$$

As a probability density function, $\phi(z)$ integrates to one which arises straightforward using the known integral value $\int_{-\infty}^{\infty} \exp(-x^2)dx = \sqrt{\pi}$ and replacing $x = z/\sqrt{2}$. The PDF of the standard Gaussian variable is shown in Figure 1.

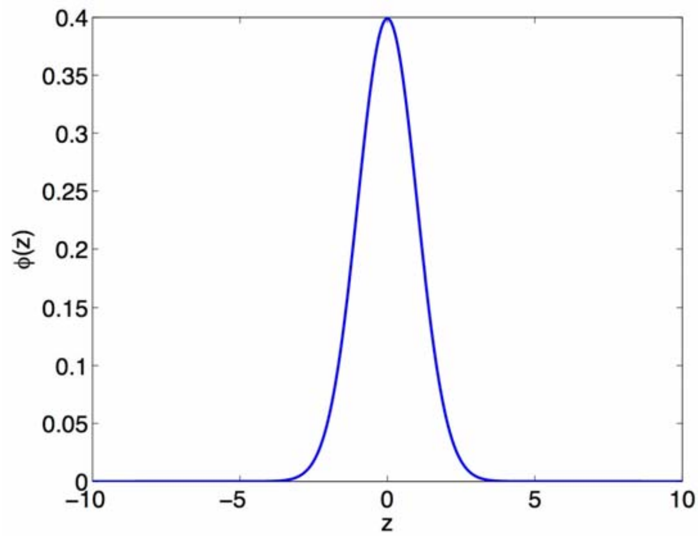

Figure 1: PDF of standard Gaussian variable

The standard normal distribution has zero mean, $\mu_Z = 0$, and unit variance, $\sigma_Z^2 = 1$. Specifically, the mean of the standard Gaussian variable Z is

$$\mu_Z = E[Z] = \int_{-\infty}^{\infty} zf(z)dz = \int_{-\infty}^{\infty} z\phi(z)dz = 0 \tag{2}$$

due to the fact that the functions $\phi(z)$ and z are even and odd functions, respectively, so that $z\phi(z)$ is an odd function and thus the integral (2) has to be zero. The variance of Z is

$$\sigma_Z^2 = E[(Z-\mu_Z)^2] = \int_{-\infty}^{\infty}(z-\mu_Z)^2\phi(z)dz = \int_{-\infty}^{\infty} z^2\phi(z)dz = 1$$

which is obtained using integration by parts, noting that $d\phi(z) = -z\phi(z)dz$ and that

$$\int_{-\infty}^{\infty} z^2\phi(z)dz = -\int_{-\infty}^{\infty} z\ d\phi(z) = -z\ \phi(z)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\phi(z)\ dz = \int_{-\infty}^{\infty}\phi(z)\ dz = 1$$

The last equality is obtained using the fact that $\phi(z)$ decays faster to zero than z does at the limits when z tends to $\pm\infty$.

Let a variable Z be a standard normal distribution and introduce the variable

$$X = \mu + \sigma Z \tag{3}$$

The variable X follows a normal distribution or Gaussian distribution with PDF

$$\begin{aligned} f(x) &= \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) \\ &= \frac{1}{\sqrt{2\pi\sigma}}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \end{aligned} \tag{4}$$

Note that the PDF in (4) is derived using the fact that for any function $X = g(Z)$ between two variables X and Z one has for the probability densities that $f_X(x)dx = f_Z(z)dz$ or

$$f_X(x) = f_Z(g^{-1}(x))\left|\left(\frac{dx}{dz}\right)^{-1}_{g^{-1}(x)}\right| \tag{5}$$

Herein, $g^{-1}(x) = (x-\mu)/\sigma$, $\dfrac{dx}{dz} = \sigma$, $f_Z(z) = \phi(z)$. Substituting in (5) one readily derives that $f_X(x) = \phi((x-\mu)/\sigma)\sigma^{-1}$ which is the same as the first of (4). The second of (4) is obtained by substituting $\phi(z)$ from (1).

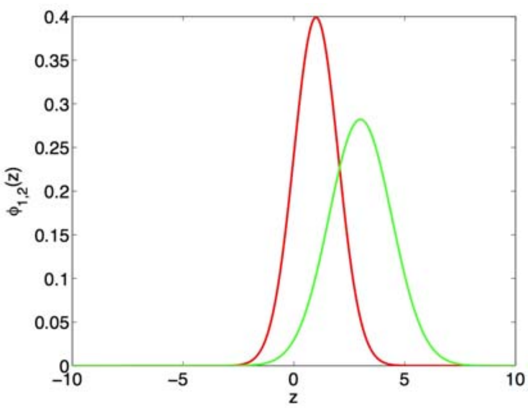

Figure 1: PDF of two Gaussian variables with mean 1 and $\sigma = 1$ (red) and mean 3 and $\sigma = 2$ (green)

The normal distribution has mean μ and variance σ^2 since $E[Z]=0$, $E[Z^2]=1$,

$$\mu_X = E[X] = E[\mu + \sigma Z] = \mu + \sigma E[Z] = \mu$$

and

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[(X-\mu)^2] = E[(\sigma Z)^2] = \sigma^2 E[Z^2] = \sigma^2$$

A Normal distribution is also denoted by $N(X;\mu,\sigma^2)$ or equivalently one can write that $X \sim N(\mu,\sigma^2)$. The parameter σ is the standard deviation and is a measure of the spread of uncertainty of the variable X around the mean value μ.

The PDF of the Gaussian variable is shown in Figure 2.

The cumulative distribution function (CDF) $F_Z(z) \equiv F(z)$ of a standard Gaussian variable or standard normal variable Z, denoted also by $\Phi(z)$, is given by the integral

$$\Phi(z) = \int_{-\infty}^{z} f_Z(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{t^2}{2}\right) dt$$

Using the error function, defined by

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp\left(-t^2\right) dt$$

the CDF of the standard Gaussian variable is given by

$$\Phi(z) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right] \text{w}$$

where the function $\Phi(z)$ satisfies $\Phi(-z) = 1 - \Phi(z)$.

The CDF of a general Gaussian variable or normal variable X is

$$F(x) = \int_{-\infty}^{x} f_X(t)dt = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) dt$$

Letting $s = (t-\mu)/\sigma$, changing the variable of integration and noting that $ds = dt/\sigma$ in the previous integrand, the CDF is given in terms of the $\Phi(x)$ as

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

1.3 Multivariate Standard Normal and General Normal Distribution

Let $\underline{Z} = (Z_1, \ldots, Z_n)^T \in R^n$ be a vector of independent and identically distributed (i.i.d.) standard Gaussian variables. Using the fact that $Z_1, \ldots, Z_n$ are independent variables, the probability density function (PDF) $f_Z(\underline{z}) \equiv f(\underline{z})$ of a standard Gaussian vector or standard normal vector $\underline{Z}$ is

$$\begin{aligned} f(\underline{z}) &= f(z_1, \ldots, z_n) = f(z_1) \ldots f(z_n) \\ &= \phi(z_1) \ldots \phi(z_n) = \frac{1}{\left(\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2}(z_1{}^2 + \ldots + z_n{}^2)\right] = \frac{1}{\left(\sqrt{2\pi}\right)^n} \exp\left[-\frac{1}{2}\underline{z}^T \underline{z}\right] \end{aligned} \quad (6)$$

The standard normal distribution has zero mean, $\underline{\mu}_Z = \underline{0}$, and identity covariance matrix, $\Sigma_Z = E[(\underline{Z} - \underline{\mu}_Z)(\underline{Z} - \underline{\mu}_Z)^T] = I$. The zero mean arises from

$$\underline{\mu}_Z = E[\underline{Z}] = \int_{-\infty}^{\infty} \underline{z} f(z) dz = 0$$

and the fact that $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} z_i f(\underline{z}) d\underline{z} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} z_i \phi(z_1) \cdots \phi(z_n) d\underline{z} = \int_{-\infty}^{\infty} z_i \phi(z_i) dz_i \prod_{\substack{j=1 \\ j \neq i}}^{n} \int_{-\infty}^{\infty} \phi(z_j) dz_j = 0$

since $\int_{-\infty}^{\infty} z_i \phi(z_i) dz_i = \mu_{Z_i} = 0$. For $i \neq j$, the (i, j) component of the covariance matrix Σ_Z is

$$\Sigma_{Z,ij} = E[(Z_i - \mu_i)(Z_j - \mu_j)] = E[(Z_i - \mu_i)]E[(Z_j - \mu_j)] = 0$$

since the variables Z_i and Z_j are independent.

For $i = j$, the (i,i) diagonal component of the covariance matrix Σ_Z is

$$\Sigma_{Z,ii} = E[(Z_i - \mu_i)^2] = \int_{-\infty}^{\infty} (z_i - \mu_i)^2 \phi(z_i) dz_i \prod_{\substack{j=1 \\ j \neq i}}^{n} \int_{-\infty}^{\infty} \phi(z_j) dz_j = 1$$

since $\int_{-\infty}^{\infty} (z_i - \mu_i)^2 \phi(z_i) dz_i = \sigma_{z_i}^2 = 1$ and $\int_{-\infty}^{\infty} \phi(z_j) dz_j = 1$ ($\phi(z_j)$ is a PDF and thus has to integrate to one).

For the special case of a bi-variate standard normal variable

$$f(z_1, z_2) = \frac{1}{2\pi} \exp\left[-\frac{1}{2}(z_1{}^2 + z_2{}^2)\right]$$

the contour plots corresponding to a level c are given by $f(z_1, z_2) = c$ or equivalently by $z_1{}^2 + z_2{}^2 = \kappa^2$, where $c = (2\pi)^{-1} \exp(-\kappa^2 / 2)$ and represent circles of radius κ in the parameter space z_1 and z_2, centered at the origin of the parameter space. The contour plots of the bi-variate standard Gaussian PDF are shown in Figure 3. The contour plots $f(z_1, z_2) = c$ quantify the spread of uncertainty of the vector $\underline{Z}$ in the parameter space (z_1, z_2).

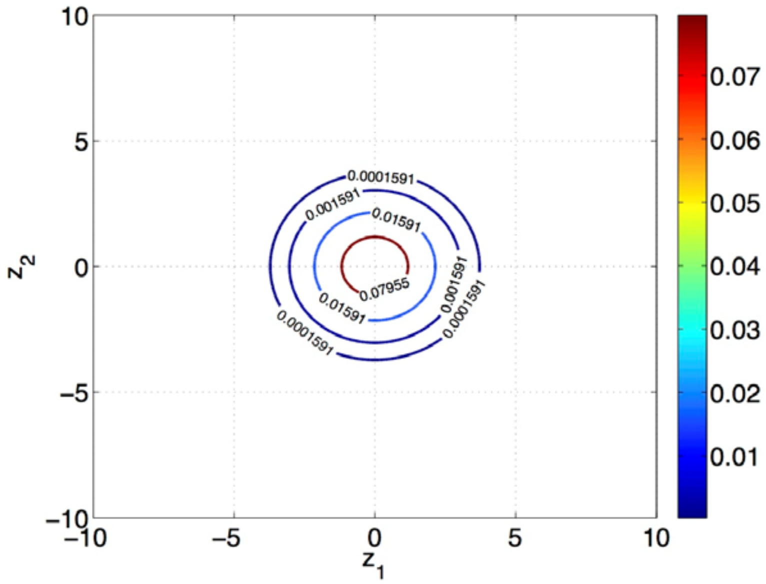

Figure 3: Contour plots $f(z_1, z_2) = c$

Let a vector $\underline{Z}$ be a multivariable standard normal distribution and introduce the vector

$$\underline{X} = \underline{\mu} + A\underline{Z} \tag{7}$$

as a linear combination of the standard normal variables in $\underline{Z}$. The vector $\underline{X}$ follows a multivariate normal distribution or multivariate Gaussian distribution with PDF

$$f(\underline{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^n |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1}(\underline{x}-\underline{\mu})\right] \tag{8}$$

where $\Sigma = AA^T$.The covariance matrix Σ of the vector $\underline{X}$ is positive semi-definite since for any vector $\underline{y} \in R^n$ with $\underline{y} \neq \underline{0}$, the quadratic form $\underline{y}^T \Sigma \underline{y} = \underline{y}^T AA^T \underline{y} = (A^T \underline{y})^T A^T \underline{y} = \underline{\xi}^T \underline{\xi} = \xi_1^2 + \ldots + \xi_n^2 \geq 0$, where $\underline{\xi} = A^T \underline{y}$. The normal distribution has mean $\underline{\mu}$ and covariance matrix Σ. This can be shown be noting that,

$$\underline{\mu}_X = E[\underline{X}] = E[\underline{\mu} + A\underline{Z}] = \underline{\mu} + AE[\underline{Z}] = \underline{\mu}$$

since $E[\underline{Z}] = \underline{0}$, and

$$\Sigma_X = E[(\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)^T] = E[(A\underline{Z})(A\underline{Z})^T] = E[A\underline{Z}\underline{Z}^T A^T] = AE[\underline{Z}\underline{Z}^T]A^T = AA^T$$

since $E[(\underline{Z} - \underline{\mu}_Z)(\underline{Z} - \underline{\mu}_Z)^T] = I$.

Note that the PDF in (8) is derived using the fact that for any function $\underline{X} = \underline{g}(\underline{Z})$ between two variables $\underline{X}$ and $\underline{Z}$ one has for the probability densities that $f_X(\underline{x})d\underline{x} = f_Z(\underline{z})d\underline{z}$ or

$$f_X(\underline{x}) = f_Z(\underline{g}^{-1}(\underline{x}))\det\left[\left(\frac{d\underline{x}}{d\underline{z}}\right)^{-1}_{\underline{g}^{-1}(\underline{x})}\right] \tag{9}$$

Herein, using (7) and (6) one has that $\underline{g}^{-1}(\underline{x}) = A^{-1}(\underline{x}-\underline{\mu})$, $\frac{d\underline{x}}{d\underline{z}} = A$, $f_Z(\underline{z}) = \frac{1}{\left(\sqrt{2\pi}\right)^n}\exp\left[-\frac{1}{2}\underline{z}^T\underline{z}\right]$.

Substituting in (9) one readily derives that

$$f_X(\underline{x}) = \frac{1}{\left(\sqrt{2\pi}\right)^n}\exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu})^T A^{-T}A^{-1}(\underline{x}-\underline{\mu})\right]|A|^{-1}$$

The Gaussian distribution (8) arises by introducing the matrix $\Sigma = AA^T$ and noting that $|\Sigma| = |A||A^T| = |A|^2$ or $|A| = |\Sigma|^{1/2}$.

A multivariate normal distribution is also denoted by $N(\underline{X};\underline{\mu},\Sigma)$ or equivalently one can write that $\underline{X} \sim N(\underline{\mu},\Sigma)$.

In order to plot the contour plots of the PDF one need to analyze the quadratic term

$$Q(\underline{x}) = (\underline{x}-\underline{\mu})^T H(\underline{x}-\underline{\mu}) \tag{10}$$

where, for convenience, it was set that $H = \Sigma^{-1}$. The contour curves $Q(\underline{x}) = \kappa$ corresponding to a level κ are exactly the same as the contour curves of $f_X(x) = c$ corresponding to the level $c = \frac{|H|^{1/2}}{\left(\sqrt{2\pi}\right)^n}\exp(-\kappa^2/2)$. The points $\underline{x}$ in the parameter space that belong to the contour curve of $Q(\underline{x})$ corresponding to a level $\kappa > 0$, have coordinates that satisfy the equation

$$Q(\underline{x}) = (\underline{x}-\underline{\mu})^T H(\underline{x}-\underline{\mu}) = \kappa^2 \tag{11}$$

In order to plot these contour curves in the two-dimensional parameter space, the following analysis and geometric interpretation of quadratic forms is required.

1.4 Quadratic Forms – Geometric Interpretation

Let the two-dimensional vectors $\underline{x} \in R^2$ and $\mu \in R^2$ with the components of these vectors defined with respect to the usual unitary basis $\{\underline{e}_1,\underline{e}_2\}$, $\underline{e}_1 = (1,0)^T \in R^2$, $\underline{e}_2 = (0,1)^T \in R^2$. Consider the eigenvalues λ_1 and λ_2 and the eigenvectors $\underline{u}_1$ and $\underline{u}_2$ of the positive definite symmetric matrix H obtained by solving the eigenvalue problem

$$H\underline{u} = \lambda\underline{u}$$

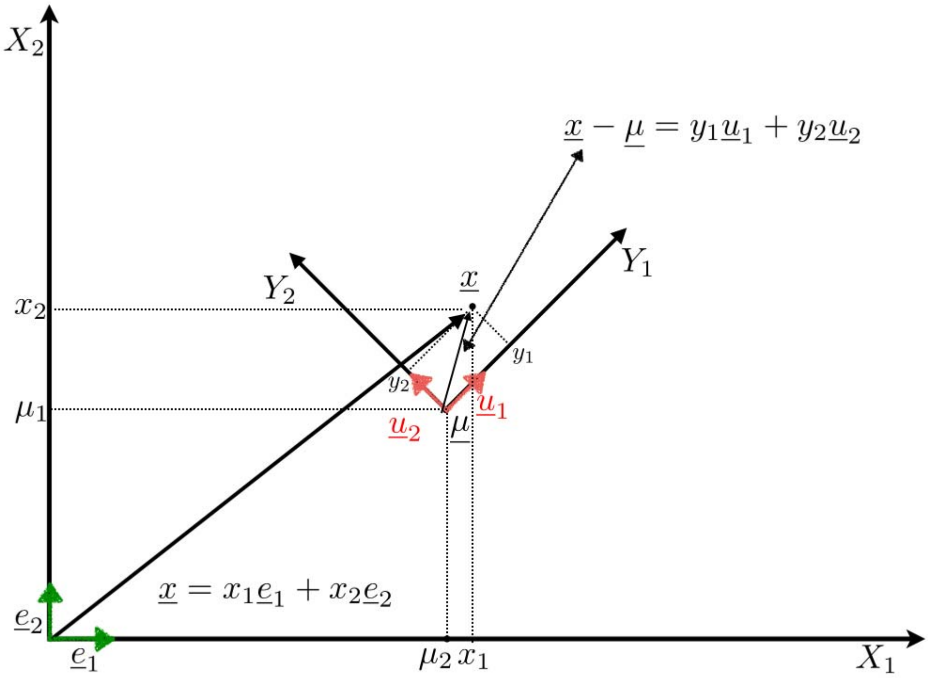

Figure 4: New coordinate system defined by the orthogonal unit vectors $\underline{u}_1$ and $\underline{u}_2$, and the coordinates y_1 and y_2 of the vector $\underline{x} - \underline{\mu}$ with respect to the new coordinate system.

From linear algebra results, it is well known that for a positive definite symmetric matrix, the eigenvalues are positive i.e. $\lambda_1 > 0$ and $\lambda_2 > 0$, while the eigenvectors $\underline{u}_1$ and $\underline{u}_2$ are orthogonal. Normalize that eigenvectors $\underline{u}_1$ and $\underline{u}_2$ so that they have unit length. These orthogonal unit vectors $\underline{u}_1$ and $\underline{u}_2$ define certain orthogonal directions in the parameter space (x_1, x_2) as shown in Figure 4. A new coordinate system is introduced, centered at the mean μ with unit vectors along the axis (y_1, y_2) of the new system to be the eigenvectors $\underline{u}_1$ and $\underline{u}_2$.

Introducing now the matrix of eigenvectors $U = [\underline{u}_1, \underline{u}_2]$ and invoking known relevant results from linear algebra, one can write the orthogonality conditions:

$$UU^T = U^T U = I$$

$$U^T H U = \Lambda$$

where Λ is the diagonal matrix of the eigenvalues of H. The first condition implies that the matrix of eigenvectors U is orthogonal. Also, from linear algebra, it is well-known that the orthonormal eigenvectors $\underline{u}_1$ and $\underline{u}_2$ constitute a basis of the two-dimensional vector space or, equivalently, any vector $\underline{x} - \underline{\mu} \in R^2$ in Figure 4 can be written in terms of the basis unit vectors $\{\underline{u}_1, \underline{u}_2\}$ in the new coordinate system as

$$\underline{x} - \underline{\mu} = y_1 \underline{u}_1 + y_2 \underline{u}_2 = [\underline{u}_1 \quad \underline{u}_2] \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = U \underline{y} \tag{12}$$

where $\underline{y} = (y_1, y_2)^T \in R^2$ are the components of the vector $\underline{x} - \underline{\mu}$ with respect to the new coordinate system defined by the orthogonal unit vectors $\{\underline{u}_1, \underline{u}_2\}$.

Substituting $\underline{x} - \underline{\mu} = U\underline{y}$ into the quadratic form (10), one derives the quadratic form $Q(x)$ in terms of the new coordinates y_1 and y_2 of the vector $\underline{x} - \underline{\mu}$ in the new coordinate system as

$$\begin{aligned} Q(\underline{x}) &= \underline{x}^T H \underline{x} = \underline{y}^T U^T H U \underline{y} = \underline{y}^T \Lambda \underline{y} = \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 = \hat{Q}(\underline{y}) \end{aligned} \qquad (13)$$

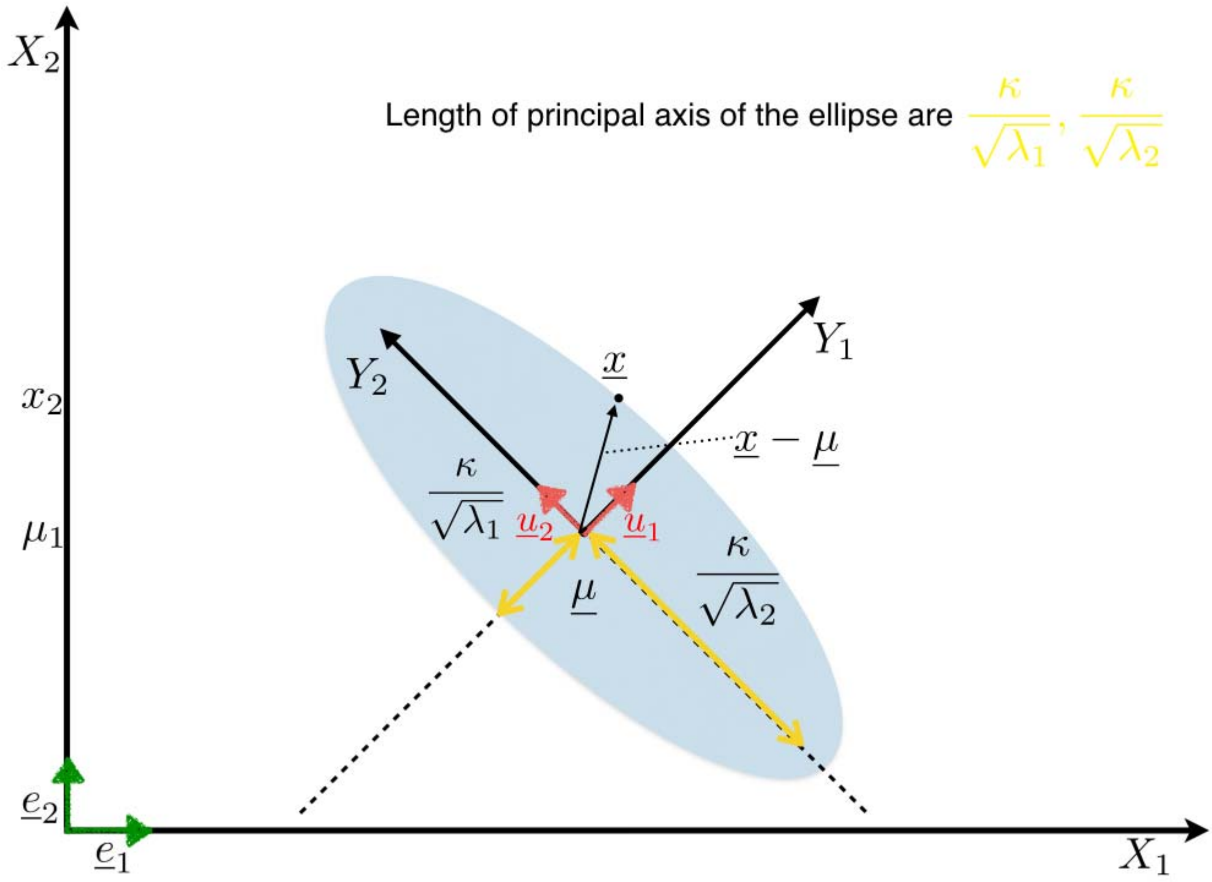

Figure 5: Contour plots $f(x_1, x_2) = c$

Consider now the points at the contour curve of the function $Q(x)$ corresponding to the "energy" level κ, satisfying the equation (11). Using (13), the points on the contour curve can conveniently be written with respect to their coordinates y_1, y_2 in the new system defined by the eigenvector basis as follows

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = \kappa^2$$

Introducing the variables $\alpha_i = \frac{\kappa}{\sqrt{\lambda_i}}$, $i = 1, 2$, this equation can be re-written in the form

$$\frac{y_1^2}{\alpha_1^2} + \frac{y_2^2}{\alpha_2^2} = 1$$

which represents an ellipse with respect to the new coordinate system (see Figure 5 for a geometric representation of the contour curves), centered at the point $\underline{\mu}$ in the parameter space with principal axis along the directions specified by the eigenvectors $\underline{u}_1$ and $\underline{u}_2$. The sizes of the principal axes of the ellipse are equal to α_1 and α_2. It is clear that the lengths of the principal axes are inversely proportional to the square root of the eigenvalues λ_1 and λ_2. Thus, the eigenvalues and the eigenvectors of the matrix H define completely the characteristics of this ellipse in the two-dimensional space. The contour plots of the bi-variate Gaussian PDF, shown in Figure 5, quantify the spread of uncertainty in the values of the parameters x_1 and x_2 in the two-dimensional parameter space (x_1, x_2) of the uncertain parameter set $\underline{X}$.

The cumulative distribution function (CDF) $F_Z(\underline{z}) \equiv F(\underline{z})$ of a standard Gaussian variable or standard normal variable $\underline{Z}$ is given by the integral

$$\begin{aligned}\Phi(\underline{z}) &= \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_n} f(\underline{t}) d\underline{t} \\ &= \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_n} \phi(t_1) \cdots \phi(t_n) dt_1 \cdots dt_n = \int_{-\infty}^{z_1} \phi(t_1) dt_1 \cdots \int_{-\infty}^{z_n} \phi(t_n) dt_n \\ &= \Phi(z_1) \cdots \Phi(z_z)\end{aligned}$$

Noting from (7) that $\underline{Z} = A^{-1}(\underline{X} - \underline{\mu})$, where A is assumed to be a non-singular matrix (non-degenerate case), and for a given value $\underline{x}$ the corresponding value of $\underline{Z}$ is $\underline{z} = A^{-1}(\underline{x} - \underline{\mu})$, the CDF of a general Gaussian vector or normal vector $\underline{X}$ is

$$\begin{aligned}F(\underline{x}) &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_X(\underline{s}) d\underline{s} \\ &= \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_n} f_Z(\underline{t}) d\underline{t} = f(\underline{z}) \\ &= \Phi(z_1) \cdots \Phi(z_n)\end{aligned}$$

where z_i is the i-th component of the vector $\underline{z} = A^{-1}(\underline{x} - \underline{\mu})$.

<u>Remarks</u>:

1. Consider the formulation for n-dimensional case.
2. Consider the formulation for the degenerate case $|\Sigma| = 0$.

1.5 Marginal of Joint Gaussian Distributions

Consider a vector $x \in R^n$ which has a Gaussian distribution with mean $\mu \in R^n$ and covariance matrix $\Sigma \in R^{n \times n}$:

$$f(\underline{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu}) \right]$$

Let a partition of the random vector $\underline{x} \in R^n$ be

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}$$

where $\underline{x}_1 \in R^{n_1}$ and $\underline{x}_2 \in R^{n_2}$, $n_1 + n_2 = 1$, are two disjoint subsets of $\underline{x}$, and let the corresponding partitions of the mean and the covariance matrix be

$$\underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

The marginal distributions of the random vector $\underline{x}_i$, $i = 1, 2$ is normal with mean $\underline{\mu}_i$ and covariance matrix Σ_{ii}, that is,

$$f(\underline{x}_i) = \frac{1}{(2\pi)^{n/2} |\Sigma_{ii}|^{1/2}} \exp\left[-\frac{1}{2} (\underline{x}_i - \underline{\mu}_i)^T \Sigma_{ii}^{-1} (\underline{x}_i - \underline{\mu}_i) \right]$$

1.6 Conditionals of Joint Gaussian Distributions

The conditional distribution of $\underline{x}_i$ given $\underline{x}_j$ ($j \neq i$) is normal with mean

$$\underline{\mu}_{i|j} = \underline{\mu}_i + \Sigma_{ij} \Sigma_{jj}^{-1} (\underline{x}_j - \underline{\mu}_j)$$

and covariance matrix

$$\Sigma_{i|j} = \Sigma_{jj} - \Sigma_{ij}^T \Sigma_{ii}^{-1} \Sigma_{ij}$$

1.7 Product of Gaussian Distributions of Same Vector Variable

The product of two Gaussian distributions $N(\underline{x}; \mu_a, \Sigma_a)$ and $N(\underline{x}; \mu_b, \Sigma_b)$ corresponding to the same vector variable $\underline{x}$ is an un-normalized Gaussian distribution given by

$$N(\underline{x}; \mu_a, \Sigma_a) N(\underline{x}; \mu_b, \Sigma_b) = z_c N(\underline{x}; \mu_c, \Sigma_c)$$

where

$$\Sigma_c = (\Sigma_a^{-1} + \Sigma_b^{-1})^{-1}$$

$$\underline{\mu}_c = \Sigma_c (\Sigma_a^{-1} \underline{\mu}_a + \Sigma_b^{-1} \underline{\mu}_b)$$

$$z_c = \frac{1}{\sqrt{2\pi}\left|\Sigma_a + \Sigma_b\right|^{1/2}} \exp\left[-\frac{1}{2}(\underline{\mu}_a - \underline{\mu}_b)^T (\Sigma_a + \Sigma_b)^{-1} (\underline{\mu}_a - \underline{\mu}_b)\right]$$

$$= \frac{1}{\sqrt{2\pi}\left|\Sigma_a \Sigma_b \Sigma_c^{-1}\right|^{1/2}} \exp\left[-\frac{1}{2}(\underline{\mu}_a - \underline{\mu}_b)^T (\Sigma_a^{-1} \Sigma_c \Sigma_b^{-1}) (\underline{\mu}_a - \underline{\mu}_b)\right]$$

1.8 Integrals of Gaussian Products

1.9 Exercises

1. The sum $Z = X + Y$ of two independent Gaussian random variables $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ is Gaussian with mean $\mu_Z = \mu_X + \mu_Y$ and variance $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$, i.e. $Z \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

   Hint: Estimate marginal distribution $f(z) = \int f(z, x)\, dx = \int f(z \mid x) f(x)\, dx$ and use the fact that $X \sim N(\mu_X, \sigma_X^2)$ and $Z \mid X \sim N(X + \mu_Y, \sigma_Y^2)$.

2. The sum $Z = X + Y$ of two Gaussian random variables $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ is Gaussian with mean $\mu_Z = \mu_X + \mu_Y$ and variance $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$, where ρ is the correlation coefficient given by $\rho = E[XY]/(\sigma_X \sigma_Y)$.

3. The mixture distribution is defined by

$$f(\underline{x}) = \sum_{i=1}^{n} w_i f_i(\underline{x})$$

   Where $f_i(\underline{x})$, $i = 1, \ldots, n$, are the mixture components and w_i are mixture weights which are non-negative $w_i \geq 0$ and satisfy $\sum_{i=1}^{n} w_i = 1$. The mixture components $f_i(\underline{x})$ are probability distributions. Show that $f(\underline{x})$ is a probability distribution. Estimate the first and second moment of the mixture distribution in terms of the first and second moment of the mixture components. Estimate the variance of the mixture distribution.

4. The mixture of Gaussian distributions is defined by

$$f(\underline{x}) = \sum_{i=1}^{n} w_i f_i(\underline{x})$$

where the mixture components $f_i(\underline{x})$, $i = 1, \ldots, n$, are Gaussian, i.e. $f_i(\underline{x}) = N(\underline{x}; \underline{\mu}_i, \Sigma_i)$ and w_i are mixture weights which are non-negative $w_i \geq 0$ and satisfy $\sum_{i=1}^{n} w_i = 1$. Estimate the mean and the variance of the mixture distribution. Find the marginal distribution of a parameter x_j in $\underline{x}$.