Gaussian Random Variables and Error Ellipse

Pierfrancesco Lombardo

Johann Carl Friedrich Gauss

Born:1777 Brunswick, Germany

Died: February 23, 1855, Göttingen, Germany

By the age of eight during arithmetic class he astonished his teachers by being able to instantly find the sum of the first hundred integers.

- Attended Brunswick College in 1792, where he discovered many important theorems before even reaching them in his studies
- Found a square root in two different ways to fifty decimal places by ingenious expansions and interpolations
- Constructed a regular 17 sided polygon, the first advance in this matter in two millennia. He was only 18 when he made the discovery

Ideas of Gauss

- Gauss was a mathematical scientist with interests in so many areas as a young man including theory of numbers, to algebra, analysis, geometry, probability, and the theory of errors.

- His interests grew, including observational astronomy, celestial mechanics, surveying, geodesy, capillarity, geomagnetism, electromagnetism, mechanism optics, and actuarial science.

- In 1805 Adrien-Marie Legendre published a paper on the method of least squares. His treatment, however, lacked a 'formal consideration of probability and it's relationship to least squares', making it impossible to determine the accuracy of the method when applied to real observations.

- Gauss claimed that he had written colleagues concerning the use of least squares dating back to 1795

Probabilistic error theory

- ***Gauss***
  - Published *'The theory of the Motion of Heavenly Bodies'* in 1809. He gave a probabilistic justification of the method,which was based on the assumption of a normal distribution of errors. Gauss himself later abandoned the use of normal error function.
  - Published *'Theory of the Combination of Observations Least Subject to Errors'* in 1820s. *He substituted the root mean square error for Laplace's mean absolute error.*

- ***Laplace*** Derived the method of least squares (between1802 and 1820) from the principle that the best estimate should have the smallest 'mean error' - the mean of the absolute value of the error.

Treatment of Errors

- Using probability theory to describe error
- Error will be treated as a random variable
- Two types of errors
  - Constant-associated with calibration
  - Random error

Error Probability Density Function

- Gauss began his study by making two assumptions:
  - Random errors of measurements of the same type lie within fixed limits
  - All errors within these limits are possible, but not necessarily with equal likelihood

We define the function $\phi(x)$ with the same meaning as a density function with the following properties.

– The probability of errors lying within the interval (x, x + dx) is $\phi(x)dx$

– Small errors are more likely to occur than large ones

– Positive and negative errors of the same maginitude are equally likely, $\phi(x) = \phi(-x)$

Mean and Variance

- Define $k = \int x\phi(x)dx$

  In many cases assume k=0

- Define mean square error as

$$m^2 = \int_{-\infty}^{\infty} x^2 \phi(x) dx$$

- If k=0 then the variance will equal m^2

Reasons for m^2

- m^2 is always positive and is simple.
- The function is differentiable and integrable unlike the absolute value function.
- The function approximates the average value in cases where large numbers of observations are being considered, and is simple to use when considering small numbers of observations.

Variance

If $k \neq 0$ then variance equals $m^2 - k^2$.

Suppose we have independent random variables $\{e, e', e'', \ldots\}$ with standard deviation 1 and expected value 0. The linear function of total errors is given by $E = \lambda e + \lambda' e' + \ldots$

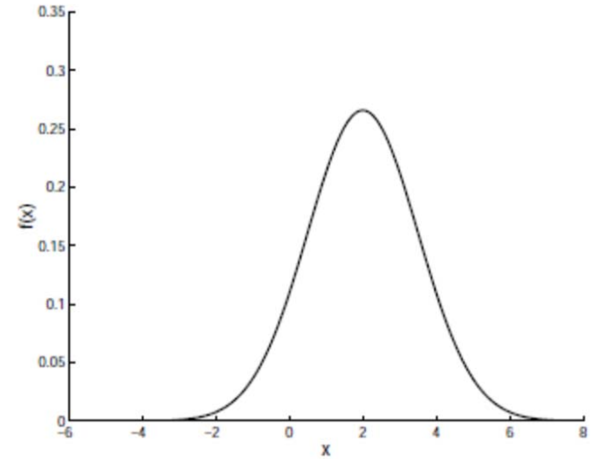
Now the variance of E is given as $M^2 = \sum_{i=1}^{k} \lambda_i^2 e_i^2 = \sum_{i=1}^{k} \lambda_i^2$

This is assuming every error falls within λ standard deviations from the mean

Gaussian pdf (I)

A random variable X is said to be normally distributed with mean μ and variance σ^2 if its probability density function (pdf) is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad -\infty < x < \infty.$$



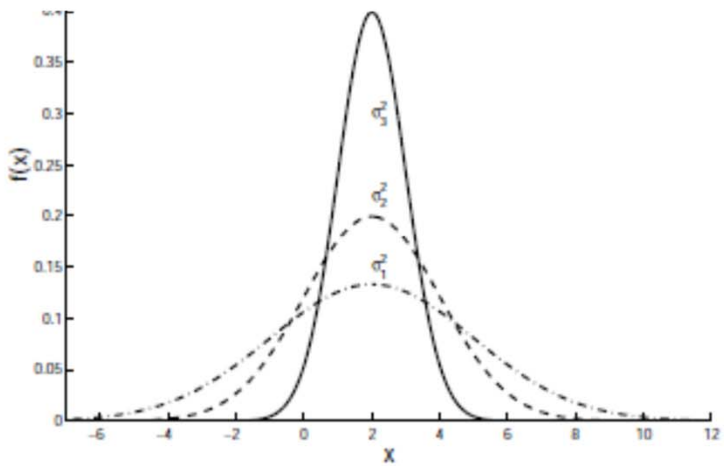
Gaussian or Normal pdf, $N(2, 1.5^2)$

The Normal or Gaussian pdf is a bell-shaped curve that is symmetric about the mean μ and that attains its maximum value of $\frac{1}{\sqrt{2\pi}\sigma} \simeq \frac{0.399}{\sigma}$ at $x = \mu$

Gaussian pdf (II)

The Gaussian pdf $\mathcal{N}(\mu, \sigma^2)$ is completely characterized by the two parameters μ and σ^2, the first and second order moments, respectively, obtainable from the pdf as

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx,$$
$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$$



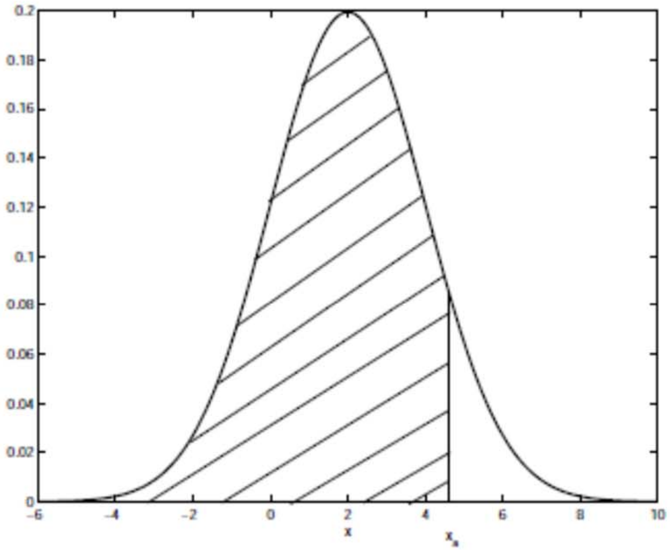
Gaussian pdf with different variances $(\sigma_1^2 = 3^2, \sigma_2^2 = 2^2, \sigma_3^2 = 1)$

Gaussian pdf (III)

Given a real number $x_a \in \mathcal{R}$, the probability that the random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ takes values less or equal x_a is given by

$$Pr\{X \leq x_a\} = \int_{-\infty}^{x_a} f(x)dx = \int_{-\infty}^{x_a} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx,$$

represented by the shaded area



Probability evaluation using pdf

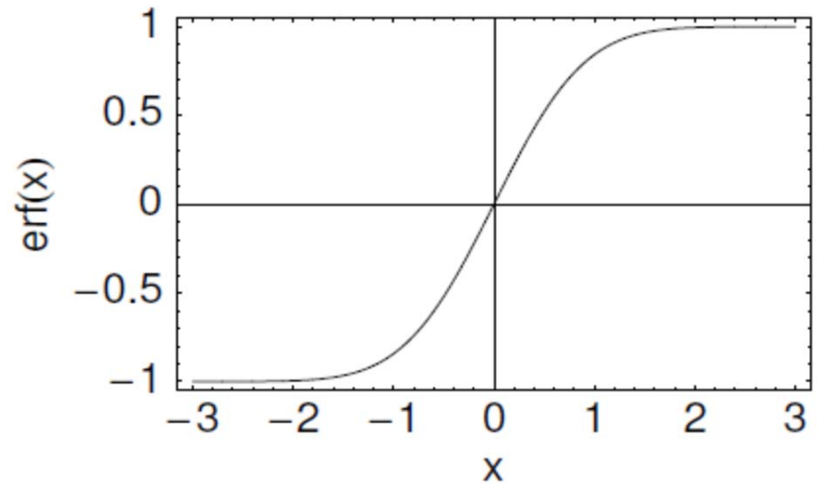
Gaussian pdf (IV)

To evaluate the probability in (1.4) the *error function*, $erf(x)$, which is related with $\mathcal{N}(0,1)$,

$$erf(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp^{-y^2/2} dy \qquad (1.5)$$

plays a key role. In fact, with a change of variables, (1.4) may be rewritte

$$Pr\{X \leq x_a\} = \begin{cases} 0.5 - erf(\frac{\mu - x_a}{\sigma}) & for \quad x_a \leq \mu \\ 0.5 + erf(\frac{x_a - \mu}{\sigma}) & for \quad x_a \geq \mu \end{cases}$$

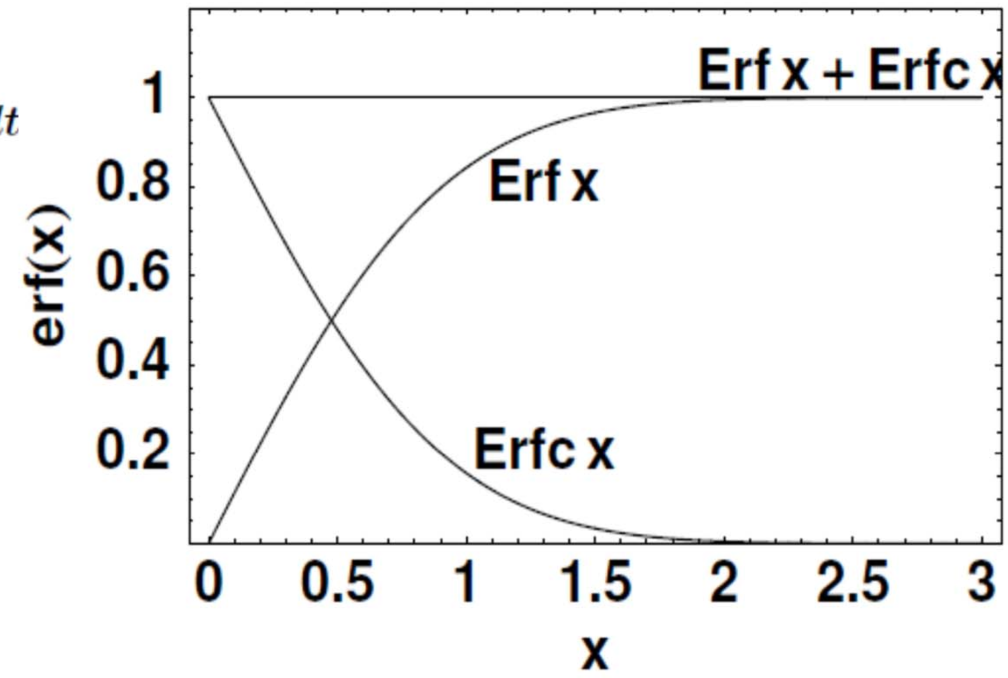


Gaussian pdf (V)

complementary Error Function

The complementary error function is defined as

$$\begin{aligned} \text{erfc}\, x &= 1 - \text{erf}\, x \\ &= \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2}\, dt \end{aligned}$$



Superposition of the Error and complementary Error Functions

Gaussian pdf (VI)

In various aspects of robotics, in particular when dealing with uncertainty in mobile robot localization, it is common the evaluation of the probability that a random variable Y (more generally a random vector representing the robot location) lies in an interval around the mean value μ. This interval is usually defined in terms of the standard deviation, σ, or its multiples.

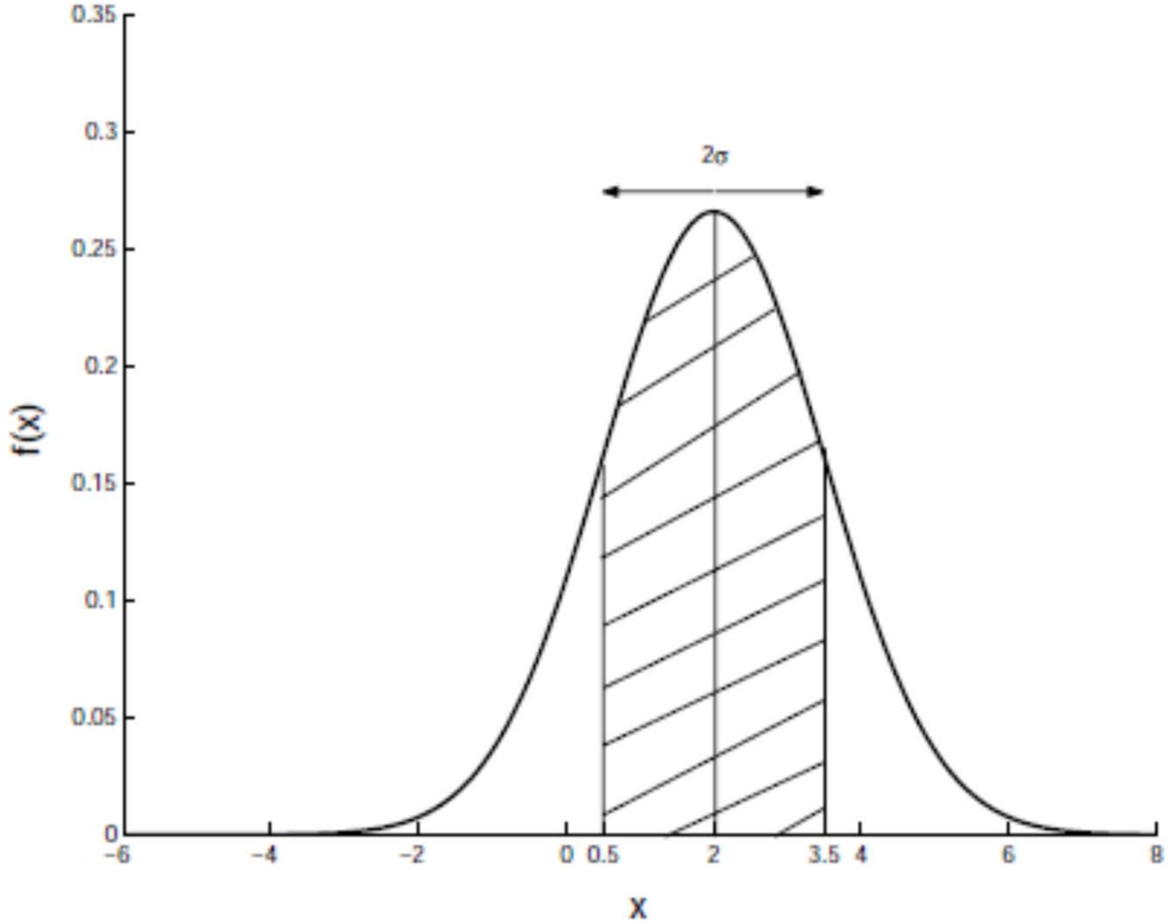
Using the error function, (1.5),the probability that the random variable X lies in an interval whose width is related with the standard deviation, is

$$\begin{aligned} Pr\{|X-\mu| \leq \sigma\} &= 2.erf(1) = 0.68268 \\ Pr\{|X-\mu| \leq 2\sigma\} &= 2.erf(2) = 0.95452 \\ Pr\{|X-\mu| \leq 3\sigma\} &= 2.erf(3) = 0.9973 \end{aligned}$$

In other words, the probability that a Gaussian random variable lies in the interval $[\mu - 3\sigma, \ \mu + 3\sigma]$ is equal to 0.9973.

Gaussian pdf (VII)

Probability of X taking values in the interval $[\mu - \sigma,\ \mu + \sigma]$, $\mu = 2, \sigma = 1.5$



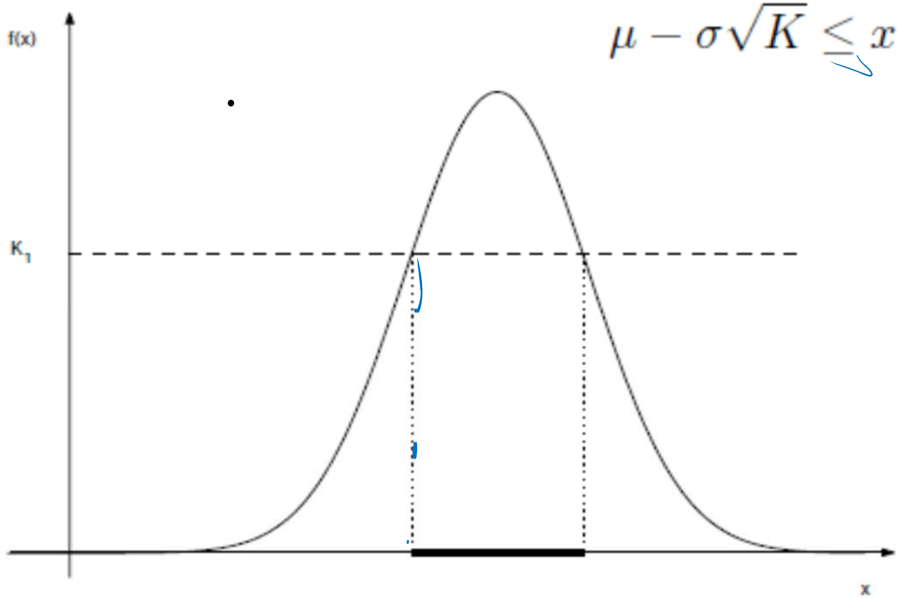
Gaussian pdf (VIII)

Another useful evaluation is the locus of values of the random variable X where the pdf is greater or equal a given pre-specified value K_1, i.e.,

$$\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \geq K_1 \iff \frac{(x-\mu)^2}{2\sigma^2} \leq K$$

with $K = -\ln(\sqrt{2\pi}\sigma K_1)$. This locus is the line segment

$$\mu - \sigma\sqrt{K} \leq x \leq \mu + \sigma\sqrt{K}$$



Locus of x where the pdf is greater or equal than K_1

Gaussian Random Vector (I)

A random vector $X = [X_1,\ X_2,\ \ldots\ X_n]^T \in \mathcal{R}^n$ is Gaussian if its pdf is

$$f_X(x) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x - m_X)^T \Sigma^{-1}(x - m_X)\right\}$$

where

- $m_X = E(X)$ is the mean vector of the random vector X,
- $\Sigma_X = E[(X - m_X)(X - m_X)^T]$ is the covariance matrix,
- $n = dimX$ is the dimension of the random vector,

also represented as

$$X \sim \mathcal{N}(m_X, \Sigma_X).$$

Gaussian Random Vector (II)

The mean vector m_X is the collection of the mean values of each of the random variables X_i,

$$m_X = E\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} m_{X_1} \\ m_{X_2} \\ \vdots \\ m_{X_n} \end{bmatrix}.$$

The covariance matrix is symmetric with elements,

$$\begin{aligned} \Sigma_X &= \Sigma_X^T = \\ &= \begin{bmatrix} E(X_1 - m_{X_1})^2 & E(X_1 - m_{X_1})(X_2 - m_{X_2}) & \dots & E(X_1 - m_{X_1})(X_n - m_{X_n}) \\ E(X_2 - m_{X_2})(X_1 - m_{X_1}) & E(X_2 - m_{X_2})^2 & \dots & E(X_2 - m_{X_2})(X_n - m_{X_n}) \\ \vdots & & & \vdots \\ E(X_n - m_{X_n})(X_1 - m_{X_1}) & \dots & \dots & E(X_n - m_{X_n})^2 \end{bmatrix}. \end{aligned}$$

The diagonal elements of Σ are the variance of the random variables X_i and the generic element $\Sigma_{ij} = E(X_i - m_{X_i})(X_j - m_{X_j})$ represents the covariance of the two random variables X_i and X_j.

Gaussian Random Vector (III)

When studying the localization of autonomous robots, the random vector X plays the role of the robot's location. Depending on the robot characteristics and on the operating environment, the location may be expressed as:

- a two-dimensional vector with the position in a 2D environment,
- a three-dimensional vector (2d-position and orientation) representing a mobile robot's location in an horizontal environment,
- a six-dimensional vector (3 positions and 3 orientations) in an underwater vehicle

When characterizing a 2D-laser scanner in a statistical framework, each range measurement is associated with a given pan angle corresponding to the scanning mechanism. Therefore the pair (distance, angle) may be considered as a random vector whose statistical characterization depends on the physical principle of the sensor device.

The above examples refer quantities, (e.g., robot position, sensor measurements) that are not deterministic. To account for the associated uncertainties, we consider them as random vectors. Moreover, we know how to deal with Gaussian random vectors that show a number of nice properties; this (but not only) pushes us to consider these random variables as been governed by a Gaussian distribution.

2D Gaussian Random Vector (I)

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \in {}^2,$$

be a second-order Gaussian random vector, with mean,

$$E[Z] = E\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} m_X \\ m_Y \end{bmatrix}$$

and covariance matrix,

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}$$

where σ_X^2 and σ_Y^2 are the variances of the random variables X and Y and σ_{XY} is the covariance of X and Y, defined below.

The covariance σ_{XY} of the two random variables X and Y is the number

$$\sigma_{XY} = E[(X - m_X)(Y - m_Y)]$$

where $m_X = E(X)$ and $m_Y = E(Y)$.

2D Gaussian Random Vector (II)

$$\begin{aligned} \sigma_{XY} &= E(XY) - m_X E(Y) - m_Y E(X) + m_X m_Y \\ &= E(XY) - E(X)E(Y) \\ &= E(X) - m_X m_Y. \end{aligned}$$

- *The* **correlation coefficient** *of the variables X and Y is defined as*

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$$|\rho| \leq 1, \qquad |\sigma_{XY}| \leq \sigma_X \sigma_Y.$$

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}.$$

For this second-order case, the Gaussian pdf particularizes as, with $z = [x \;\; y]^T \in \mathcal{R}^2$,

$$\begin{aligned} f(z) &= \frac{1}{2\pi\sqrt{det\Sigma}} \exp\left[-\frac{1}{2}[x - m_X \; y - m_Y]\Sigma^{-1}[x - m_X \; y - m_Y]^T\right] \\ &= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-m_X)^2}{\sigma_X^2} - \frac{2\rho(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y} + \frac{(y-m_Y)^2}{\sigma_Y^2}\right)\right] \end{aligned}$$

2D Gaussian Random Vector (III)

Two random variables X and Y are called independent if the joint pdf, $f(x, y)$ equals the product of the pdf of each random variable, $f(x)$, $f(y)$, i.e.,

$$f(x, y) = f(x)f(y)$$

In the case of Gaussian random variables, clearly X and Y are independent when $\rho = 0$. This issue will be further explored later.

Two random varibles X and Y are called uncorrelated if their covariance is zero, i.e.,

$$\sigma_{XY} = E[(X - m_X)(Y - m_Y)] = 0,$$

which can be written in the following equivalent forms:

$$\rho = 0, \quad E(XY) = E(X)E(Y).$$

If two random variables X and Y are independent, then they are uncorrelated, i.e.,

$$f(x, y) = f(x)f(y) \Rightarrow E(XY) = E(X)E(Y)$$

but the converse is not, in general, true.

2D Gaussian Random Vector (IV)

Variance of the sum of two random variables *Let* X *and* Y *be two random variables, jointly distributed, with mean* m_X *and* m_Y *and correlation coefficient* ρ *and let*

$$Z = X + Y.$$

Then,

$$E(Z) = m_Z = E(X) + E(Y) = m_X + m_Y$$
$$\sigma_Z^2 = E[(Z - m_Z)^2)] = \sigma_X^2 + 2\rho\sigma_X\sigma_Y + \sigma_Y^2.$$

Variance of the sum of two uncorrelated random variables

Let X *and* Y *be two uncorrelated random variables, jointly distributed, with mean* m_X *and* m_Y *and let*

$$Z = X + Y.$$

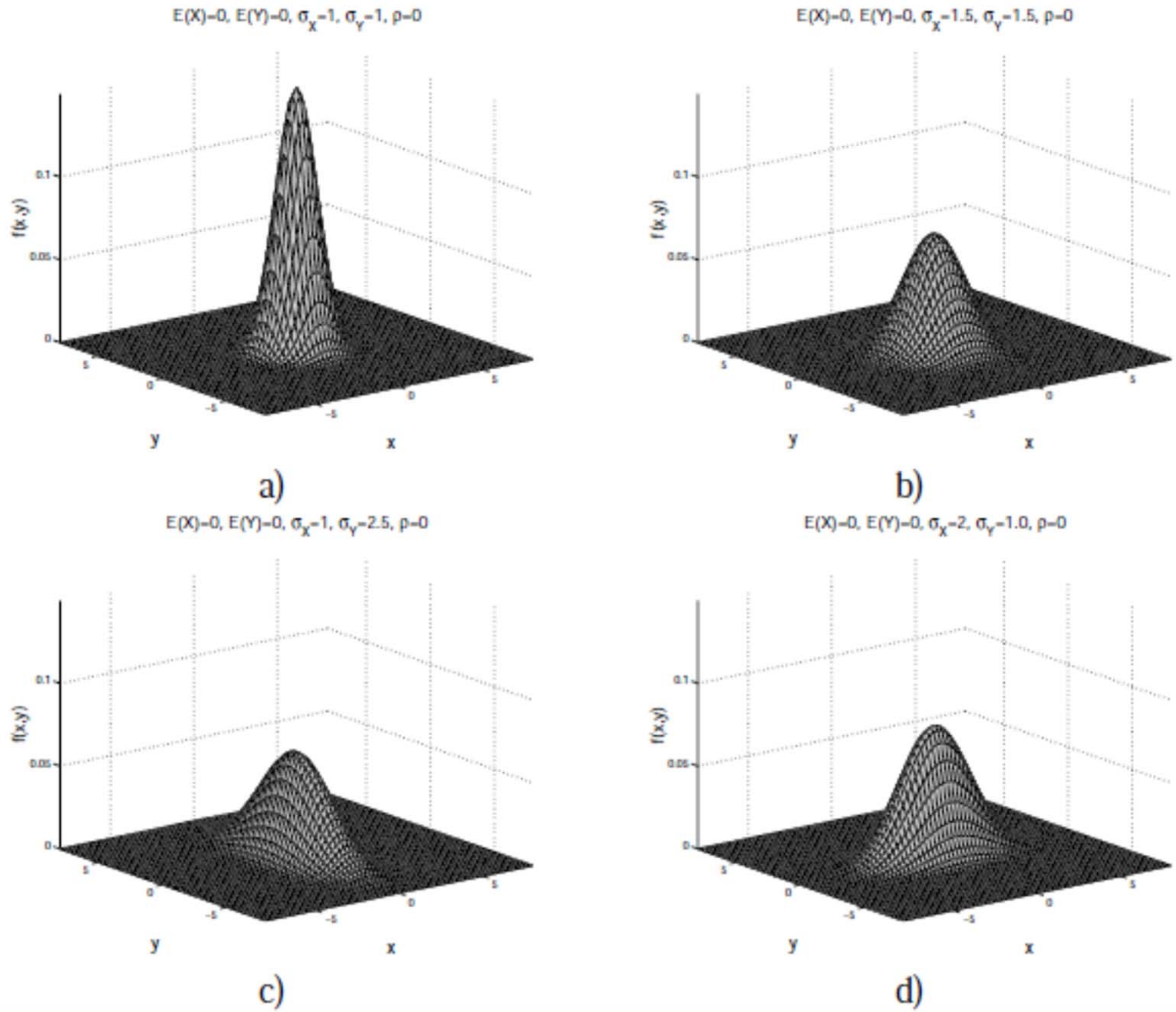
Then,

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$$

i.e., if two random variables are uncorrelated, then the variance of their sum equals the sum of their variances.

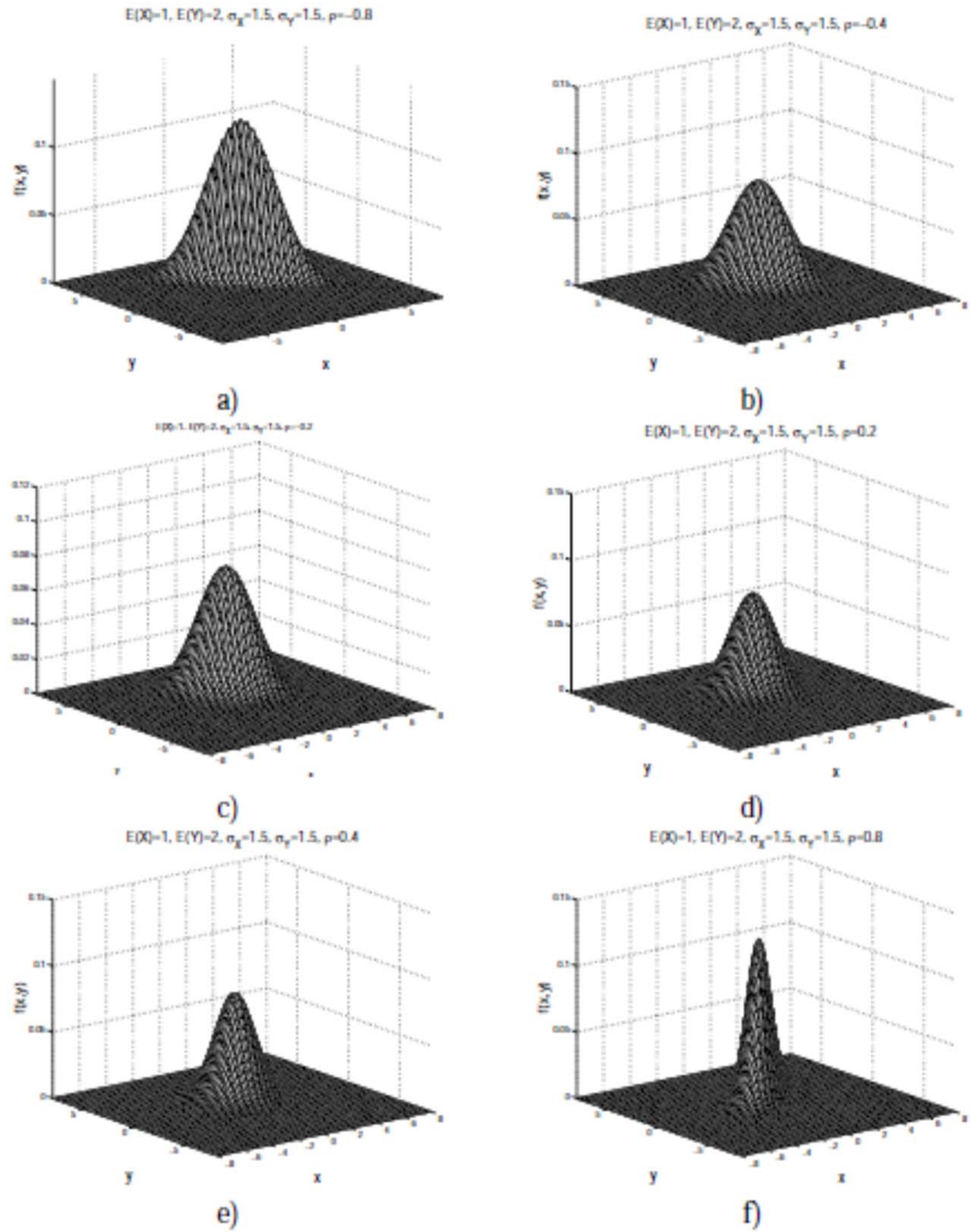
2D Gaussian Random Vector (V)

$\rho = 0$



2D Gaussian Random Vector (VI)

$$\rho = \begin{matrix} -0{,}8 \\ -0{,}4 \\ -0{,}2 \\ 0{,}2 \\ 0{,}4 \\ 0{,}8 \end{matrix}$$



2D Gaussian Random Vector (VII)

Locus of constant probability

Similarly to what was considered for a Gaussian random variable, it is also useful for a variety of applications and for a second order Gaussian random vector, to evaluate the locus (x, y) for which the pdf is greater or equal a specified constant, K_1, i.e.,

$$\left\{(x,y) : \frac{1}{2\pi\sqrt{det\Sigma}} \exp\left[-\frac{1}{2}[x - m_X \; y - m_Y]\Sigma^{-1}[x - m_X \; y - m_Y]^T\right] \geq K_1\right\}$$

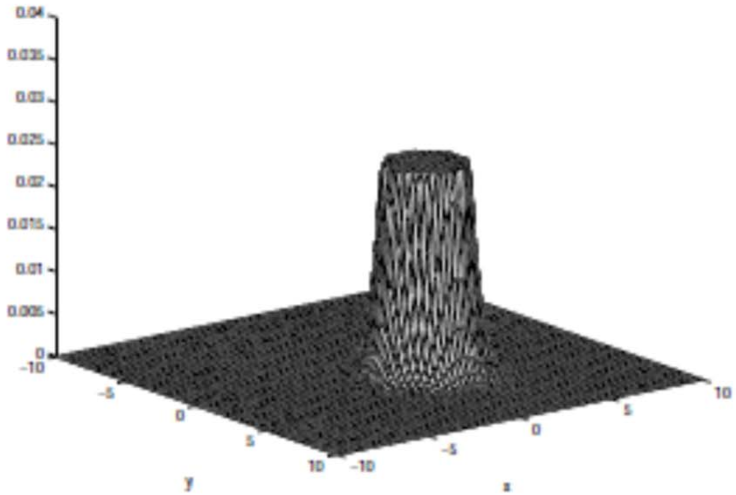
which is equivalent to

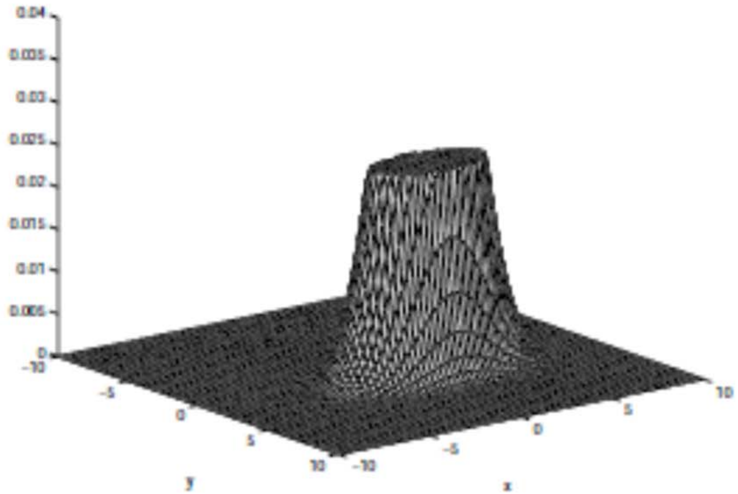
$$\left\{(x,y) : [x - m_X \; y - m_Y]\Sigma^{-1}\begin{bmatrix} x - m_X \\ y - m_Y \end{bmatrix} \leq K\right\}$$

with

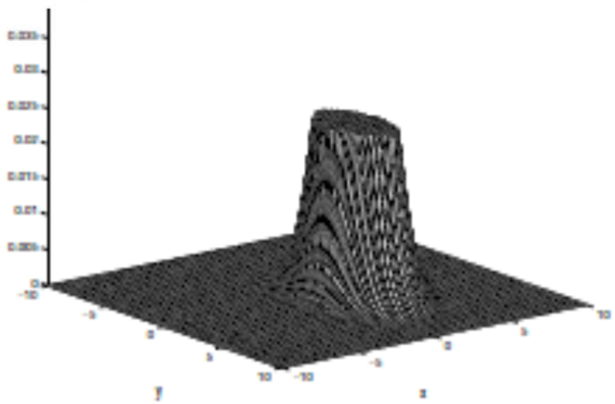
$$K = -2\ln(2\pi K_1\sqrt{det\Sigma}).$$

2D Gaussian Random Vector (VIII)

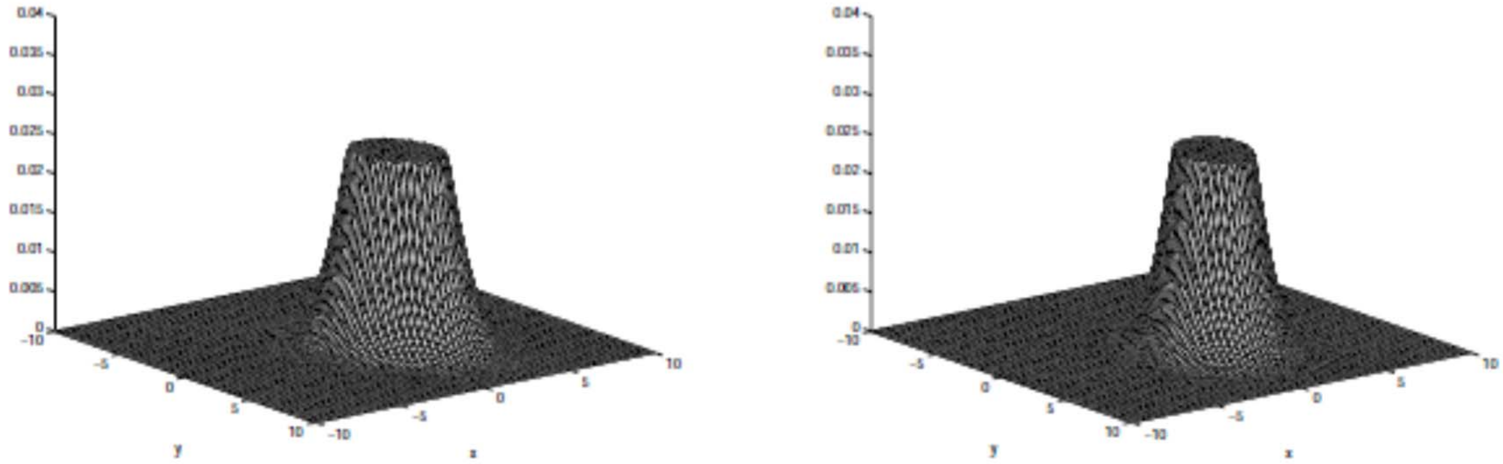




Locus of (x, y) such that the Gaussian pdf is less than a constant K for uncorrelated Gaussian random variables with $m_X = 1, \ m_Y = 2$ - a) $\sigma_X = \sigma_Y = 1$, b) $\sigma_X = 2, \ \sigma_Y = 1$, c) $\sigma_X = 1, \ \sigma_Y = 2$



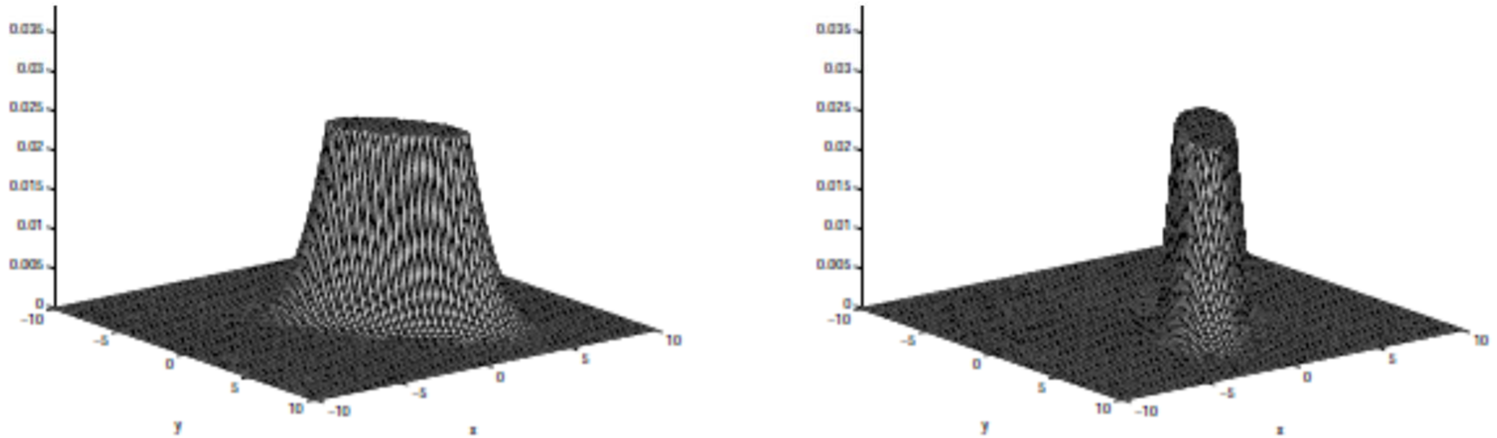
2D Gaussian Random Vector (IX)



Locus of (x, y) such that the Gaussian pdf is less than a constant K for correlated variables . $m_X = 1,\ m_Y = 2,\ \sigma_X = 1.5,\ \sigma_Y = 2$ - a) $\rho = 0.2$, b) $\rho = 0.7$, c) $\rho = -0.2$, d) $\rho = -0.7$

c)

E(X)=1, E(Y)=2, σ_x=1.5, σ_y=2, ρ=-0.7



2D Gaussian Random Vector (X)

The locus in (2.16) is the border and the inner points of an ellipse, centered in $(m_X,\ m_Y)$. The length of the ellipses axis and the angle they do with the axis x and y are a function of the constant K, of the eigenvalues of the covariance matrix

Σ and of the correlation coefficient. We will demonstrate this statement in two different steps. We show that:

1. **Case 1** - if Σ in (2.16) is a diagonal matrix, which happens when $\rho = 0$, i.e., X and Y are uncorrelated, the ellipse axis are parallel to the frame axis.
2. **Case 2** - if Σ in (2.16) is non-diagonal, i.e, $\rho \neq 0$, the ellipse axis are not parallel to the frame axis.

In both cases, the length of the ellipse axis is related with the eigenvalues of the covariance matrix Σ in (2.3) given by:

$$\lambda_1 = \frac{1}{2}\left[\sigma_X^2 + \sigma_Y^2 + \sqrt{(\sigma_X^2 - \sigma_Y^2)^2 + 4\sigma_X^2\sigma_Y^2\rho^2}\right], \qquad (2.17)$$

$$\lambda_2 = \frac{1}{2}\left[\sigma_X^2 + \sigma_Y^2 - \sqrt{(\sigma_X^2 - \sigma_Y^2)^2 + 4\sigma_X^2\sigma_Y^2\rho^2}\right]. \qquad (2.18)$$

2D Gaussian Random Vector (XI)

Case 1 - Diagonal covariance matrix

When $\rho = 0$, i.e., the variables X and Y are uncorrelated, the covariance matrix is diagonal,

$$\Sigma = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

and the eigenvalues particularize to $\lambda_1 = \sigma_X^2$ and $\lambda_2 = \sigma_Y^2$. In this particular case, illustrated in Figure 2.5, the locus (2.16) may be written as

$$\left\{(x, y) : \frac{(x - m_X)^2}{\sigma_X^2} + \frac{(y - m_Y)^2}{\sigma_Y^2} \leq K\right\}$$

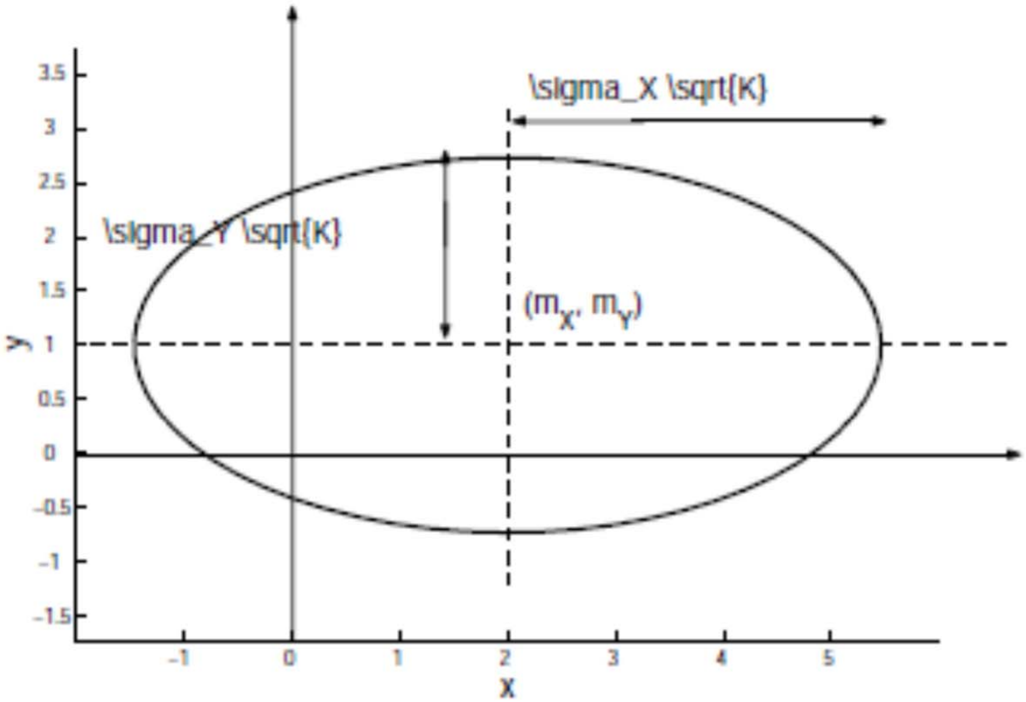
or also,

$$\left\{(x, y) : \frac{(x - m_X)^2}{K\sigma_X^2} + \frac{(y - m_Y)^2}{K\sigma_Y^2} \leq 1\right\}.$$

2D Gaussian Random Vector (XII)

ellipse that is the border of the locus

- x-axis with length $2\sigma_X \sqrt{K}$
- y-axis with length $2\sigma_Y \sqrt{K}$.



Locus of constant pdf: ellipse with axis parallel to the frame axis

2D Gaussian Random Vector (XIII)

Case 2 - Non-diagonal covariance matrix

When the covariance matrix Σ is non-diagonal, the ellipse that borders the locus has center in (m_X, m_Y) but its axis are not aligned with the coordinate frame. In the sequel we evaluate the angle between the ellipse axis and those of the coordinate frame. With no loss of generality we will consider that $m_X = m_Y = 0$, i.e., the ellipse is centered in the coordinated frame. Therefore, the locus under analysis is given by

$$\left\{(x, y) : [x \;\; y]\Sigma^{-1}\begin{bmatrix} x \\ y \end{bmatrix} \leq K\right\}$$

where Σ is the matrix As it is a symmetric matrix, the eigenvectors corresponding to distinct eigenvalues are orthogonal.

2D Gaussian Random Vector (XIV)

When $\sigma_X \neq \sigma_Y$

the eigenvalues are distinct, the corresponding eigenvectors are orthogonal and therefore Σ has simple structure which means that there exists a non-singular and unitary coordinate transformation T such that

$$\Sigma = T\,D\,T^{-1}$$

where

$$T = [\,v_1 \quad | \quad v_2\,], \quad D = diag(\lambda_1,\ \lambda_2)$$

and v_1, v_2 are the unit-norm eigenvectors of Σ associated with λ_1 and λ_2.

$$\left\{(x,y) : [x \quad y] T D^{-1} T^{-1} \begin{bmatrix} x \\ y \end{bmatrix} \leq K \right\}.$$

2D Gaussian Random Vector (XV)

Denoting

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = T^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$$

and given that $T^T = T^{-1}$, it is immediate that

$$\left\{ (w_1, w_2) : [w_1 \quad w_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \leq K \right\}$$

that corresponds, in the new coordinate system defined by the axis w_1 and w_2, to the locus bordered by an ellipse aligned with those axis. Given that v_1 and v_2 are unit-norm orthogonal vectors, the coordinate transformation defined corresponds to a rotation of the coordinate system (x, y), around its origin by an angle

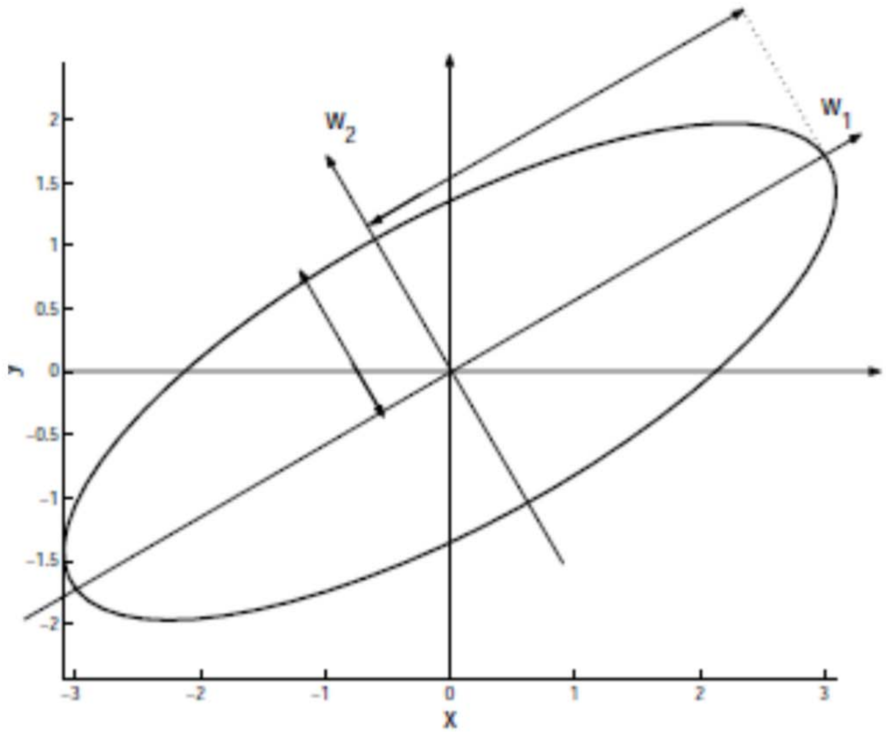
$$\alpha = \frac{1}{2} \tan^{-1} \left(\frac{2\rho\sigma_X\sigma_Y}{\sigma_X^2 - \sigma_Y^2} \right), \quad -\frac{\pi}{4} \leq \alpha \leq \frac{\pi}{4}, \quad \sigma_X \neq \sigma_Y.$$

2D Gaussian Random Vector (XVI)

$$\left\{(w_1, w_2) : \frac{w_1^2}{K\lambda_1} + \frac{w_2^2}{K\lambda_2} \leq 1\right\}$$

that corresponds to an ellipse having

- w_1-axis with length $2\sqrt{K\lambda_1}$
- w_2-axis with length $2\sqrt{K\lambda_2}$



Ellipses non-aligned with the coordinate axis x and y.

Error Ellipsoid (I)

Let X be a n-dimensional Gaussian random vector, with

$$X \sim \mathcal{N}(m_X, \Sigma_X)$$

and consider a constant, $K_1 \in \mathcal{R}$. The locus for which the pdf $f(x)$ is greater or equal a specified constant K_1,i.e.,

$$\left\{ x : \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}[x - m_X]^T \Sigma_X^{-1} [x - m_X] \right] \geq K_1 \right\} \tag{4.1}$$

which is equivalent to

$$\left\{ x : [x - m_X]^T \Sigma_X^{-1} [x - m_X] \leq K \right\} \tag{4.2}$$

with $K = -2\ln((2\pi)^{n/2} K_1 |\Sigma|\ 1/2)$ is an n-dimensional ellipsoid centered at the mean m_X and whose axis are only aligned with the cartesian frame if the covariance matrix Σ is diagonal. The ellipsoid defined by (4.2) is the region of minimum volume that contains a given probability mass under the Gaussian assumption.

Error Ellipsoid (II)

When in (4.2) rather than having an inequality there is an equality,

$$\left\{x : [x - m_X]^T \Sigma_X^{-1} [x - m_X] = K\right\}$$

this locus may be interpreted as the contours of equal probability.

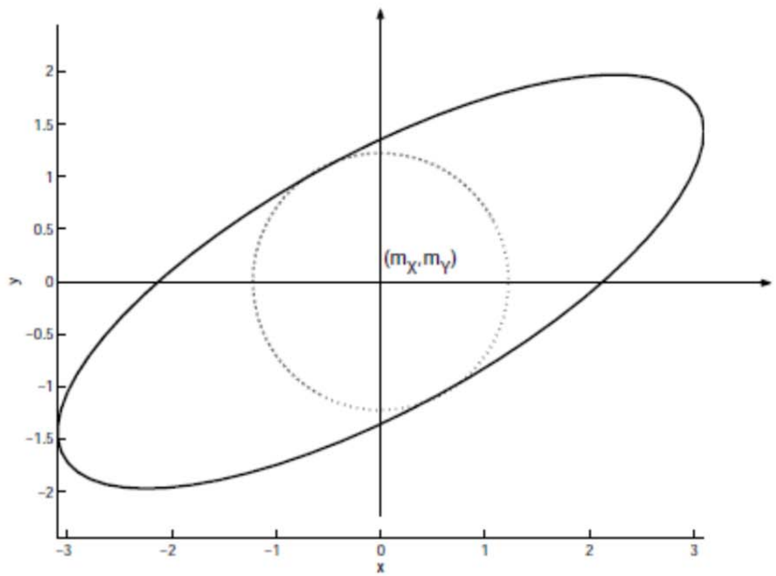
Definition 4.1 Mahalanobis distance *The scalar quantity*

$$[x - m_X]^T \Sigma_X^{-1} [x - m_X] = K$$

is known as the Mahalanobis distance of the vector x to the mean m_X.

Error Ellipsoid (III)

The Mahalanobis distance, is a normalized distance where normalization is achieved through the covariance matrix. The surfaces on which K is constant are ellipsoids that are centered about the mean m_X, and whose semi-axis are $\sqrt{K}$ times the eigenvalues of Σ_X, as seen before. In the special case where the random variables that are the components of X are uncorrelated and with the same variance, i.e., the covariance matrix Σ is a diagonal matrix with all its diagonal elements equal, these surfaces are spheres, and the Mahalanobis distance becomes equivalent to the Euclidean distance.



Error Ellipsoid (IV)

For decision making purposes (e.g., the field-of-view, a validation gate), and given m_X and Σ_X, it is necessary to determine the probability that a given vector will lie within, say, the 90% confidence ellipse or ellipsoid given by (4.3). For a given K, the relationship between K and the probability of lying within the ellipsoid specified by K is, [3],

$$\begin{aligned} n=1;\quad & Pr\{\text{x inside the ellipsoid}\} = -\tfrac{1}{\sqrt{2\pi}} + 2erf(\sqrt{K}) \\ n=2;\quad & Pr\{\text{xinside the ellipsoid}\} = 1 - e^{-K/2} \\ n=3;\quad & Pr\{\text{x inside the ellipsoid}\} = -\tfrac{1}{\sqrt{2\pi}} + 2erf(\sqrt{K}) - \sqrt{\tfrac{2}{\pi}}\sqrt{K}e^{-K/2} \end{aligned} \tag{4.4}$$

where n is the dimension of the random vector. Numeric values of the above expression for $n = 2$ are presented in the following table

Probability	K
50%	1.386
60%	1.832
70%	2.408
80%	3.219
90%	4.605

For a given K the ellispoid axis are fixed. The probability that a given value of the random vector X lies within the ellipsoid centered in the mean value, increases with the increase of K.

Derivation of Error Probability (I)

Result 4.1 *Given the n-dimensional Gaussian random vector X, with mean m_X and covariance matrix Σ_X, the scalar random variable K defined by the quadratic form*

$$[x - m_X]^T \Sigma_X^{-1} [x - m_X] = K \qquad (4.5)$$

has a chi-square distribution with n degrees of freedom.

Proof: see, p.e., in [1].

The pdf of K in (4.5), i.e., the chi-square density with n degrees of freedom is, (see, p.e., [1])

$$f(k) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} k^{\frac{n-2}{2}} \exp^{-\frac{k}{2}}$$

where the gamma function satisfies,

$$\Gamma(\frac{1}{2}) = \sqrt{\pi}, \quad \Gamma(1) = 1 \quad \Gamma(n+1) = \Gamma(n).$$

Derivation of Error Probability (II)

The probability that the scalar random variable, K in (4.5) is less or equal a given constant, χ_p^2

$$Pr\{K \leq \chi_p^2\} = Pr\{[x - m_X]^T \Sigma^{-1} [x - m_X] \leq \chi_p^2\} = p$$

is given in the following table where n is the number of degrees of freedom and the sub-indice p in χ_p^2 represents the corresponding probability under evaluation.

n	$\chi^2_{0.995}$	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.75}$	$\chi^2_{0.50}$	$\chi^2_{0.25}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$
1	7.88	6.63	5.02	3.84	2.71	1.32	0.455	0.102	0.0158	0.0039
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	0.575	0.211	0.103
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	0.584	0.352
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	0.711

From this table we can conclude, for example, that for a third-order Gaussian random vector, $n = 3$,

$$Pr\{K \leq 6.25\} = Pr\{[x - m_X]^T \Sigma^{-1} [x - m_X] \leq 6.25\} = 0.9$$

Example (I)

Consider a mobile platform, moving in an environment and let $P \in \mathcal{R}^2$ be the position of the platform relative to a world frame. P has two components,

$$P = \begin{bmatrix} X \\ Y \end{bmatrix}.$$

The exact value of P is not known and we have to use any particular localization algorithm to evaluate P. The most common algorithms combine internal and external sensor measurements to yield an estimated value of P.

The uncertainty associated with all the quantities involved in this procedure, namely vehicle model, sensor measurements, environment map representation, leads to consider P as a random vector. Gaussianity is assumed for simplicity.

Therefore, the localization algorithm provides an estimated value of P, denoted as $\hat{P}$, which is the mean value of the Gaussian pdf, and the associated covariance matrix, i.e.,

$$P \sim \mathcal{N}(\hat{P}, \Sigma_P)$$

Example (II)

At each time step of the algorithm we do not know the exact value of P, *but we have an estimated value,* $\hat{P}$ *and a measure of the uncertainty of this estimate, given by* Σ_P. *The evident question is the following: "Where is the robot?", i.e., "What is the exact value of* P*"? It is not possible to give a direct answer to this question, but rather a probabilistic one. We may answer, for example: "Given* $\hat{P}$ *and* Σ_P, *with 90% of probability, the robot is located in an ellipse centered in* $\hat{P}$ *and whose border is defined according to the Mahalanobis distance". In this case the value of* K *in (4.5) will be* $K = 4.61$.

Someone may say that, for the involved application, a probability of 90% *is small and ask to have an answer with an higher probability, for example* 99%. *The answer will be similar but, in this case, the error ellipse, will be defined for* $K = 9.21$, *i.e.,the ellipse will be larger than the previous one.*