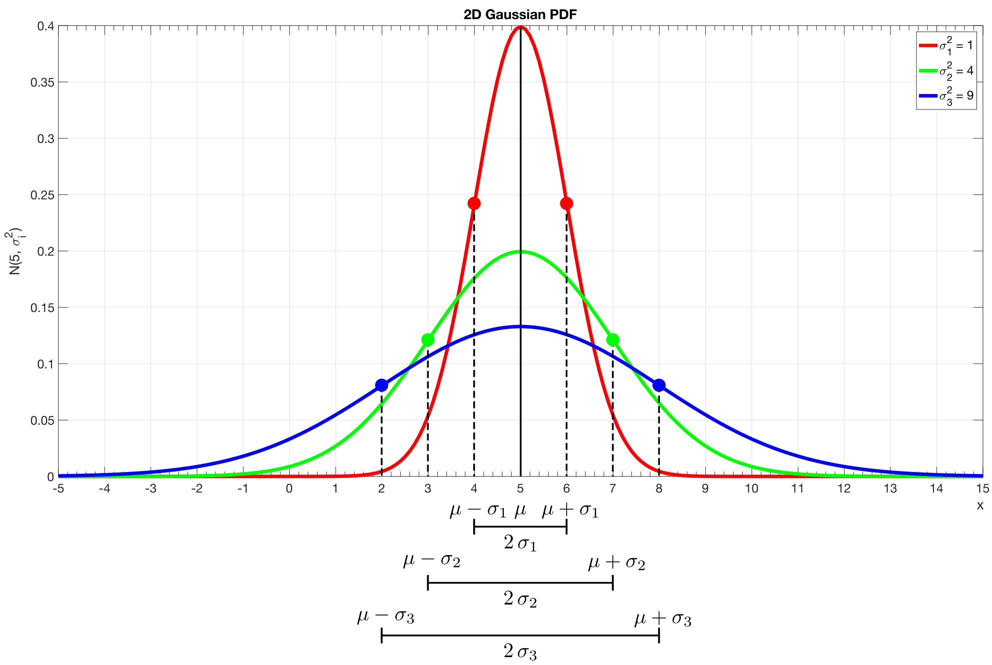
L04: THE KALMAN FILTER

THE UNIVARIATE AND MULTIVARIATE GAUSSIAN PROBABILITY DENSITY FUNCTION.

Univariate Gaussian probability density function.
The variable x has dimension 1.

$$p\,(x) \;=\; \mathcal{N}\left(\mu, \sigma^2\right) \;=\; \frac{1}{\sqrt{2\,\pi}\,\sigma} \cdot e^{\frac{-(x\,-\,\mu)^2}{2\,\sigma^2}}$$



The term μ is called mean,. The mean value is the value of the random variable where the Gaussian PDF has its maximum value and it's also the value around which the remaining values of that random variable are distributed.

The quadratic form $\left(\frac{x-\mu}{\sigma}\right)^2$ represents a straight segment centered at the mean that contains all the values of the random variable x that gives a value greater or equal to K_0 in the Gaussian PDF.

$$\frac{1}{\sqrt{2\,\pi}\,\sigma} \cdot e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} \geq K_0$$

$$\begin{aligned}
e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} &\geq K_0 \cdot \sqrt{2\,\pi}\,\sigma \\
-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2 &\geq ln\left(K_0 \cdot \sqrt{2\,\pi}\,\sigma\right) \\
\left(\frac{x-\mu}{\sigma}\right)^2 &\leq -2 \cdot ln\left(K_0 \cdot \sqrt{2\,\pi}\,\sigma\right) \\
\left(\frac{x-\mu}{\sigma}\right)^2 &\leq ln\left(K_0 \cdot \sqrt{2\,\pi}\,\sigma\right)^{-2} \\
\left(\frac{x-\mu}{\sigma}\right)^2 &\leq ln\left(\frac{1}{\left(K_0 \cdot \sqrt{2\,\pi}\,\sigma\right)^2}\right) \\
\left(\frac{x-\mu}{\sigma}\right)^2 &\leq ln\left(\frac{1}{K_0^2 \cdot 2\,\pi\,\sigma^2}\right)
\end{aligned}$$

Let's call K to the term $ln\left(\frac{1}{K_0^2 \cdot 2\,\pi\,\sigma^2}\right)$ for simplicity.

$$\begin{aligned}
\left(\frac{x-\mu}{\sigma}\right)^2 &\leq K \\
(x-\mu)^2 &\leq K\,\sigma^2 \\
\pm\,(x-\mu) &\leq K\,\sigma
\end{aligned}$$

$$\mu - K\,\sigma \leq x \leq \mu + K\,\sigma \longrightarrow \text{size} = 2\,K\,\sigma$$

$$Pr\,(a \leq x \leq b) = \int_a^b \mathcal{N}\left(\mu, \sigma^2\right) = \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_a^b e^{\frac{-(x-\mu)^2}{2\sigma^2}}\,dx$$

The error function is defined as:

$$\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-t^2}\,dt = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$$

$$
\begin{aligned}
Pr\,(a \le x \le b) &= \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_a^b e^{\frac{-(x-\mu)^2}{2\sigma^2}}\,dx = \\
&= \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_a^b e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx = \\
&= \frac{1}{\sqrt{2\,\pi}\,\sigma}\left(\int_0^b e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx - \int_0^a e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx\right) = \\
&= \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_0^b e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx - \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_0^a e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx
\end{aligned}
$$

Let’s change the integration variable:

$$
\begin{aligned}
t &= \left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right) \\
dt &= \frac{1}{\sqrt{2}\,\sigma}\,dx \\
dx &= \sqrt{2}\,\sigma\,dt
\end{aligned}
$$

$$
\begin{aligned}
x = b &\longrightarrow t = \left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right) \\
x = a &\longrightarrow t = \left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)
\end{aligned}
$$

$$
\begin{aligned}
Pr\,(a \le x \le b) &= \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_a^b e^{\frac{-(x-\mu)^2}{2\sigma^2}}\,dx = \\
&= \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_0^b e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx - \frac{1}{\sqrt{2\,\pi}\,\sigma}\int_0^a e^{-\left(\frac{x-\mu}{\sqrt{2}\,\sigma}\right)^2}\,dx = \\
&= \frac{\sqrt{2}\,\sigma}{\sqrt{2\,\pi}\,\sigma}\int_0^{\left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt - \frac{\sqrt{2}\,\sigma}{\sqrt{2\,\pi}\,\sigma}\int_0^{\left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dx = \\
&= \frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt - \frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt &= \frac{2}{2}\,\frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt = \frac{1}{2}\,\mathrm{erf}\left(\frac{b-\mu}{\sqrt{2}\,\sigma}\right) \\
\frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt &= \frac{2}{2}\,\frac{1}{\sqrt{\pi}}\int_0^{\left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)} e^{-t^2}\,dt = \frac{1}{2}\,\mathrm{erf}\left(\frac{a-\mu}{\sqrt{2}\,\sigma}\right)
\end{aligned}
$$

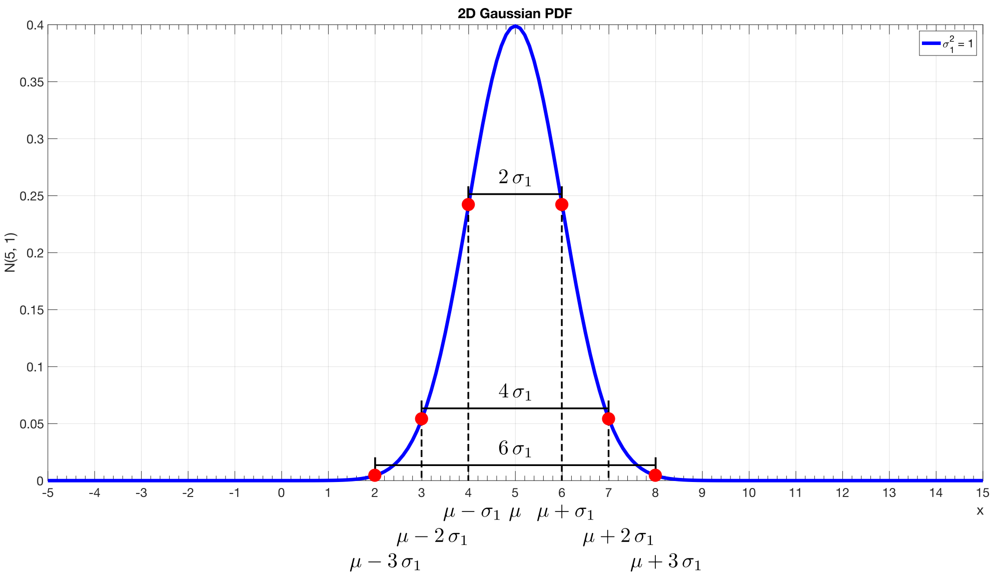
$$
Pr\,(a \le x \le b) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{b-\mu}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{a-\mu}{\sqrt{2}\sigma}\right)\right]
$$

$$Pr\,(\mu\,-\,K\,\sigma\,\leq\,x\,\leq\,\mu\,+\,K\,\sigma)\,=\,\frac{1}{2}\left[\mathrm{erf}\left(\frac{K}{\sqrt{2}}\right)-\mathrm{erf}\left(\frac{-K}{\sqrt{2}}\right)\right]$$

$$\mathrm{erf}\,(-x)\,=\,-\,\mathrm{erf}\,(x)$$

$$Pr\,(\mu\,-\,K\,\sigma\,\leq\,x\,\leq\,\mu\,+\,K\,\sigma)\,=\,\mathrm{erf}\left(\frac{K}{\sqrt{2}}\right)$$

$$K\,=\,1\;\longrightarrow\;Pr\,(\mu\,-\,1\,\sigma\,\leq\,x\,\leq\,\mu\,+\,1\,\sigma)\,=\,\mathrm{erf}\left(\frac{1}{\sqrt{2}}\right)\,=\,0.682689\;\;(68.269\%)$$
$$K\,=\,2\;\longrightarrow\;Pr\,(\mu\,-\,2\,\sigma\,\leq\,x\,\leq\,\mu\,+\,2\,\sigma)\,=\,\mathrm{erf}\left(\frac{2}{\sqrt{2}}\right)\,=\,0.954499\;\;(95.450\%)$$
$$K\,=\,3\;\longrightarrow\;Pr\,(\mu\,-\,3\,\sigma\,\leq\,x\,\leq\,\mu\,+\,3\,\sigma)\,=\,\mathrm{erf}\left(\frac{3}{\sqrt{2}}\right)\,=\,0.997300\;\;(99.730\%)$$



Multivariate Gaussian probability density function.

The term $\vec{x}$ is a real vector with N dimensions.

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_N \end{pmatrix} \in \mathbb{R}^{N \times 1}$$

:

$$p(\vec{x}) = \mathcal{N}(\vec{\mu}, \Sigma) = \frac{1}{(2\,\pi)^{N/2}\,\left|\Sigma\right|^{1/2}} \cdot e^{-\frac{1}{2} \cdot (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})}$$

where the normalizing constant $\frac{1}{(2\,\pi)^{N/2}\,\left|\Sigma\right|^{1/2}}$ makes the volume under the multivariate Gaussian PDF equal to 1.

$$0 \leq \mathcal{N}(\vec{\mu}, \Sigma) \leq \infty$$

$$0 \leq \frac{1}{(2\,\pi)^{N/2}\,\left|\Sigma\right|^{1/2}} \cdot e^{-\frac{1}{2} \cdot (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})} \leq \infty$$

The vector $\vec{\mu}$ is called mean. The mean vector $\vec{\mu}$ is a real vector with N dimensions. The mean vector $\vec{\mu}$ is the value of the random vector $\vec{x}$ where the Multivariate Gaussian PDF has its maximum value and it's also the value around which the remaining values of that random vector $\vec{x}$ are distributed.

$$\vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_j \\ \vdots \\ \mu_N \end{pmatrix} \in \mathbb{R}^{N \times 1}$$

The matrix Σ is called covariance matrix. The covariance matrix Σ is a real symmetric matrix with N dimensions. $\Sigma \in \mathbb{R}^{N \times N}$. Obviously, all the variances and covariances of the covariance matrix Σ are positive:

$$\sigma_i^2 \geq 0, \forall\, i = \{1, \dots N\}$$
$$\sigma_{ij} \geq 0 \;\forall\, i = \{1, \dots N\}, \forall\, j = \{1, \dots N\},\, i \neq j$$

Note that we haven't defined what are the relations among theses variances and covariances yet. We will back to this issue later when we have explained something else.

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1x_2} & \dots & \sigma_{x_1x_j} & \dots & \sigma_{x_1x_N} \\ \sigma_{x_1x_2} & \sigma_{x_2}^2 & \dots & \sigma_{x_2x_j} & \dots & \sigma_{x_2x_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \sigma_{x_1x_j} & \sigma_{x_2x_j} & \dots & \sigma_{x_j}^2 & \dots & \sigma_{x_jx_N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_1x_N} & \sigma_{x_2x_N} & \dots & \sigma_{x_jx_N} & \dots & \sigma_{x_N}^2 \end{pmatrix}$$

Let's talk about the covariance matrix Σ. Since the covariance matrix Σ is symmetric and real **the spectral descomposition theorem** stands that:
Any real symmetric matrix Σ, with dimensions $N \times N$, can be written as:

$$\Sigma = U \cdot D \cdot U^T$$

where:

- The matrix D is a real diagonal matrix, with dimensions $N \times N$, where the entries of the matrix D are the eigenvalues of the matrix Σ.

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_j & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & \lambda_N \end{pmatrix}$$

  Note: The matrix Σ has N eigenvalues. Not all the N eigenvalues are necessarily unique. Some of them might be repeated. It's fine.

- The matrix U is a real orthonormal matrix, with dimensions $N \times N$, whose columns are the eigenvectors of the matrix Σ.

$$U = \begin{pmatrix} | & | & & | & & | \\ \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_j & \dots & \vec{u}_N \\ | & | & & | & & | \end{pmatrix}$$

$$U^T = U^{-1} \longrightarrow U^T \cdot U = I \longrightarrow \begin{cases} \langle \vec{u}_j, \vec{u}_k \rangle = \vec{u}_j^T \cdot \vec{u}_k = 0, \text{ for } j \neq k \\ \langle \vec{u}_j, \vec{u}_k \rangle = \vec{u}_j^T \cdot \vec{u}_k = 1, \text{ for } j = k \end{cases}$$

$$\Sigma \cdot \vec{u}_j = \lambda_j \cdot \vec{u}_j, \; \forall \; j = \{1, \ldots, N\}$$

Now, that it has been shown how the matrix D and the matrix U are we can also express the matrix Σ as:

$$\begin{aligned}
&\Sigma = U \cdot D \cdot U^T = \\
&= \begin{pmatrix} | & | & & | & & | \\ \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_j & \ldots & \vec{u}_N \\ | & | & & | & & | \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_j & & \\ & & & & \ddots & \\ & & & & & \lambda_N \end{pmatrix} \cdot \begin{pmatrix} — & \vec{u}_1 & — \\ — & \vec{u}_2 & — \\ & \vdots & \\ — & \vec{u}_j & — \\ & \vdots & \\ — & \vec{u}_N & — \end{pmatrix} = \\
&= \lambda_1 \vec{u}_1 \cdot \vec{u}_1^T + \lambda_2 \vec{u}_2 \cdot \vec{u}_2^T + \ldots \lambda_j \vec{u}_j \cdot \vec{u}_j^T + \ldots + \lambda_N \vec{u}_N \cdot \vec{u}_N^T = \\
&= \sum_{j=1}^{N} \lambda_j \vec{u}_j \vec{u}_j^T
\end{aligned}$$

Let's show the relation between the covariance matrix Σ and its inverse Σ^{-1}:

$$\Sigma^{-1} = \left(U \cdot D \cdot U^T\right)^{-1} = \left(U^T\right)^{-1} \cdot D^{-1} \cdot U^{-1} = \left(U^{-1}\right)^{-1} \cdot D^{-1} \cdot U^T = U \cdot D^{-1} \cdot U^T$$

Therefore, the matrix Σ^{-1}, symmetric and real, with dimensions $N \times N$, can also be described in terms of eigenvalues and enigenvectors. The eigenvectors of the matrix Σ^{-1} are the eigenvectors of the matrix Σ. The eigenvalues of the matrix Σ^{-1} are the inverse of the eigenvalues of the matrix Σ. Because of the matrix D is diagonal it's very easy to calculate it's inverse.

$$D^{-1} = \begin{pmatrix} 1/\lambda_1 & & & & & \\ & 1/\lambda_2 & & & & \\ & & \ddots & & & \\ & & & 1/\lambda_j & & \\ & & & & \ddots & \\ & & & & & 1/\lambda_N \end{pmatrix}$$

$$\Sigma^{-1} = U \cdot D^{-1} \cdot U^T = \sum_{j=1}^{N} \frac{1}{\lambda_j} \vec{u}_j \vec{u}_j^T$$

The spectral theorem implies that there is a transformation from the real symmetric matrix Σ^{-1} into a real diagonal matrix D^{-1}. Let's explain why the transformation from a real symmetric matrix into a real diagonal

matrix is important.
Every quadratic function with N variables can be expressed as:

$$\begin{aligned} q(\vec{y}) &= \langle \vec{y},\ A \cdot \vec{y} \rangle = \\ &= \vec{y}^T \cdot A \cdot \vec{y} = \\ &= \sum_{j=1}^{N} \sum_{k=1}^{N} A_{jk} \cdot y_j \cdot y_k \end{aligned}$$

The formula for $q(\vec{y})$ involves N^2 terms, and the variables are typically coupled. However if the matrix A happens to be a diagonal matrix, then the formula for $q(\vec{y})$ simplifies considerably:

$$q(\vec{y}) = \sum_{j=1}^{N} A_{jj} y_j^2$$

Such a quadratic form is easy to understand: In each coordinate direction y_j the graph is a parabola, opening upward if $A_{jj} > 0$ and opening downward if $A_{jj} < 0$. There is also the degenerate case $A_{jj} = 0$, in which case q is constant with respect to y_j and the graph in that direction is a horizontal line.
The quadratic form

$$\left\langle (\vec{x} - \vec{\mu}),\ \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right\rangle = (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})$$

is the squared statistical distance from the vector $\vec{\mu}$ to the vector $\vec{x}$. This squared statistical distance takes into account the variances and covariances among the components of the random vector $\vec{x}$. The squared statistical distance $(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})$ is called the **squared Mahalanobis distance**. The Mahalanobis distance reduces to the Euclidean distance when the matrix Σ is the identity matrix.
Let's do a change of variable:

$$\vec{y} = (\vec{x} - \vec{\mu})$$

$$\begin{aligned} \left\langle (\vec{x} - \vec{\mu}),\ \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right\rangle &= (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) = \\ &= \vec{y}^T \cdot \Sigma^{-1} \cdot \vec{y} = \\ &= \vec{y}^T \cdot U \cdot D^{-1} \cdot U^T \cdot \vec{y} = \\ &= \left(U^T \cdot \vec{y}\right)^T \cdot D^{-1} \cdot \left(U^T \cdot \vec{y}\right) = \\ &= \left\langle \left(U^T \cdot \vec{y}\right),\ D^{-1} \cdot \left(U^T \cdot \vec{y}\right) \right\rangle \end{aligned}$$

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_N \end{pmatrix} = U^T \cdot \vec{y} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1k} & \dots & u_{1N} \\ u_{21} & u_{22} & \dots & u_{2k} & \dots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ u_{j1} & u_{j2} & \dots & u_{jk} & \dots & u_{jN} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{N1} & u_{N2} & \dots & u_{Nk} & \dots & u_{NN} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_j \\ \vdots \\ y_N \end{pmatrix} =$$

$$= \begin{pmatrix} \vec{u}_1^T,\ \vec{y} \\ \vec{u}_2^T,\ \vec{y} \\ \vdots \\ \vec{u}_j^T,\ \vec{y} \\ \vdots \\ \vec{u}_N^T,\ \vec{y} \end{pmatrix} = \begin{pmatrix} \langle \vec{u}_1,\ \vec{y} \rangle \\ \langle \vec{u}_2,\ \vec{y} \rangle \\ \vdots \\ \langle \vec{u}_j,\ \vec{y} \rangle \\ \vdots \\ \langle \vec{u}_N,\ \vec{y} \rangle \end{pmatrix}$$

$$\left(U^T \cdot \vec{y}\right)^T \cdot D^{-1} \cdot \left(U^T \cdot \vec{y}\right)$$

$$\left(\alpha_1, \dots, \alpha_j, \dots, \alpha_N\right) \cdot \begin{pmatrix} 1/\lambda_1 & & & & \\ & \ddots & & & \\ & & 1/\lambda_j & & \\ & & & \ddots & \\ & & & & 1/\lambda_N \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_N \end{pmatrix}$$

$$\left(\tfrac{1}{\lambda_1}\,\alpha_1, \dots, \tfrac{1}{\lambda_j}\,\alpha_j, \dots, \tfrac{1}{\lambda_N}\alpha_N\right) \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_N \end{pmatrix}$$

$$\frac{1}{\lambda_1}\,\alpha_1^2 + \dots + \frac{1}{\lambda_j}\,\alpha_j^2 + \dots + \frac{1}{\lambda_N}\alpha_N^2 = \sum_{j=1}^{N} \frac{1}{\lambda_j} \cdot \alpha_j^2$$

$$\left\langle (\vec{x} - \vec{\mu}),\, \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right\rangle = (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) = \sum_{j=1}^{N} \frac{1}{\lambda_j} \cdot \alpha_j^2 =$$
$$= \sum_{j=1}^{N} \frac{1}{\lambda_j} \cdot \langle \vec{u}_j, (\vec{x} - \vec{\mu}) \rangle^2 =$$
$$= \sum_{j=1}^{N} \frac{1}{\lambda_j} \cdot \left(\vec{u}_j^T \cdot (\vec{x} - \vec{\mu}) \right)^2$$

The term $\alpha_j = \langle \vec{u}_j, (\vec{x} - \vec{\mu}) \rangle = \vec{u}_j^T \cdot (\vec{x} - \vec{\mu})$ represents the $j - th$ coordinate of the vector $(\vec{x} - \vec{\mu})$ in the orthonormal basis U.
Since every real symmetric matrix A has a spectral decomposition in real eigenvalues and real eigenvectors, this means that every quadratic function $q(\vec{y}) = \left(\vec{y} - \vec{b}\right)^T \cdot A \cdot (\vec{y} - b)$ can be expressed as above
The set of all the values of the random vector $\vec{x}$ that gives a constant value K^2 in the quadratic function $(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})$ forms a N-dimensional curve. Let's call this set S. These N-dimensional curves also give a constant value L on the Gaussian PDF, that is why these curves are called contour curves.

$$q(\vec{x}) = (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \mu) = \sum_{j=1}^{N} \frac{1}{\lambda_j} \cdot \left(\vec{u}_j^T \cdot (\vec{x} - \vec{\mu}) \right)^2 = K^2$$
$$S = \{\vec{x}_1, \vec{x}_2, \ldots\} \text{ forms a N-dimensional curve that solves the above constant expression}$$
$$p(\vec{x}) = \frac{1}{(2\pi)^{N/2} \left|\Sigma\right|^{1/2}} \cdot e^{-\frac{1}{2} K^2} = L$$

For the Gaussian PDF to be well defined, it is necessary for all the eigenvalues λ_j of the covariance matrix Σ to be strictly positive, otherwise the Gaussian PDF cannot be properly normalized. If all of the eigenvalues λ_j are strictly positive, then these N-dimensional curves have the shape of N-dimensional ellipsoids with their centres at the vector $\vec{\mu}$, the direction of its $j - th$ major axis defined by the eigenvector $\vec{u}_j$ of the covariance matrix Σ and with the length of its $j - th$ major axis defined by $2K\sqrt{\lambda_j}$.
Summarizing, the covariance matrix Σ is a real symmetric matrix, with dimensions $N \times N$ that can be written in terms of its N eigenvalues and its N orthonormal eigenvectors using the expression

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} & \cdots & \sigma_{x_1 x_j} & \cdots & \sigma_{x_1 x_N} \\ \sigma_{x_1 x_2} & \sigma_{x_2}^2 & \cdots & \sigma_{x_2 x_j} & \cdots & \sigma_{x_2 x_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \sigma_{x_1 x_j} & \sigma_{x_2 x_j} & \cdots & \sigma_{x_j}^2 & \cdots & \sigma_{x_j x_N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_1 x_N} & \sigma_{x_2 x_N} & \cdots & \sigma_{x_j x_N} & \cdots & \sigma_{x_N}^2 \end{pmatrix} = U \cdot D \cdot U^T$$

and also all its eigenvalues are strictly positive, $\lambda_j > 0, \forall j = 1, \ldots, N$
Any real symmetric matrix A with all its eigenvalues strictly positive is called **positive definite**. Positive definite matrices are very important in all kind of contexts for its unique properties. In any case, I'm not

going to write about this properties here because it's not part the topic of this document.
Therefore, the covariance matrix Σ is a positive definite matrix. So, imagine your are writing a real symmetric matrix A and you want to know if this real symmetric matrix has all its eigenvalues strictly positive without having to calculate all of them. Well, there is a test you can do. The test consist on looking at the N upper left determinants of the matrix A, i.e, $\left|A_j\right|$, where A_j is the upper left $j \times j$ submatrix. All the eigenvalues of our real symmetric matrix A will be strictly positive if and only if $\left|A_j\right| > 0, \ \forall \ 1 \leq j \leq N$
Is the following matrix positive definite?

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\begin{aligned}
\left|A_1\right| &= 2 > 0 \\
\left|A_2\right| &= \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 2 \cdot 2 - (-1 \cdot -1) = 4 - 1 = 3 > 0 \\
\left|A_3\right| &= \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} = (2 \cdot 2 \cdot 2) + (-1 \cdot -1 \cdot 0) + (0 \cdot -1 \cdot -1) - \\
&- \Big((0 \cdot 2 \cdot 0) + (-1 \cdot -1 \cdot 2) + (2 \cdot -1 \cdot -1)\Big) = 8 - \Big(2 + 2\Big) = 8 - 4 = 4 > 0
\end{aligned}$$

$\left|A_j\right| > 0, \ \forall \ 1 \leq j \leq 3$, so the matrix A is positive definite.
So, for the matrix Σ, these constraints of having $\left|\Sigma_j\right| > 0, \ \forall \ 1 \leq j \leq N$ establish some relationships among the variances and covariances that constitute the matrix Σ.
Note: the constraints and conditions that apply to the matrix Σ are the same that the ones that apply to the matrix Σ^{-1}. Therefore, we don't have the necessity of analyzing two different sets of constraints and conditions.
For example:

$$N = 2 \longrightarrow \Sigma = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma_{x_2}^2 \end{pmatrix}$$

$$\left|\Sigma_1\right| = \sigma_{x_2}^2 > 0 \longrightarrow \left|\sigma_{x_2}\right| > 0$$

$$\left|\sigma_{x_2}\right| > 0 \longrightarrow \begin{cases} \sigma_{x_2} > 0 \longrightarrow \left|\sigma_{x_2}\right| = \sigma_{x_2} > 0 & \longrightarrow \sigma_{x_2} > 0 \\ \sigma_{x_2} < 0 \longrightarrow \left|\sigma_{x_2}\right| = -\sigma_{x_2} > 0 & \longrightarrow \sigma_{x_2} < 0 \end{cases}$$

The condition $\sigma_{x_2} < 0$ is not valid since the standard deviation is a measure of the dispersion of the data, so it must be positive, not negative.
When analyzing the $\left|\sigma_{x_2}\right|$ I haven't taken into account the value $\sigma_{x_2} = 0$, only the value $\sigma_{x_2} > 0$, because, although a standard deviation can be 0, this value express that I am working with a deterministic value, not a random variable. Because of I'm working with random variables (packed in the form of a vector), i.e, non deterministic values, the value $\sigma_{x_2} = 0$ is not considered.

$$\left|\Sigma_2\right| = \sigma_{x_2}^2 \cdot \sigma_{x_1}^2 - (\sigma_{x_1x_2} \cdot \sigma_{x_1x_2}) = \sigma_{x_2}^2 \cdot \sigma_{x_1}^2 - \sigma_{x_1x_2}^2 > 0$$

$$\sigma_{x_2}^2 \cdot \sigma_{x_1}^2 - \sigma_{x_1x_2}^2 > 0 \longrightarrow (\sigma_{x_2} \cdot \sigma_{x_1})^2 > \sigma_{x_1x_2}^2 \longrightarrow \left|\sigma_{x_2} \cdot \sigma_{x_1}\right| > \left|\sigma_{x_1x_2}\right|$$

A standard deviation is a measure of the dispersion of the data, therefore, it must be positive:

$$\sigma_j > 0, \; \forall \; j = 1, \ldots, N$$

(for random variables we only use the symbol >)

$$\sigma_{x_1} > 0$$

$$\sigma_{x_2} > 0$$

(equal symbol is supressed, due to the above condition)

$$\sigma_{x_1} \cdot \sigma_{x_2} > 0 \longrightarrow \left|\sigma_{x_1} \cdot \sigma_{x_2}\right| = \sigma_{x_1} \cdot \sigma_{x_2}$$

$$\left|\sigma_{x_2} \cdot \sigma_{x_1}\right| > \left|\sigma_{x_1x_2}\right| \longrightarrow \sigma_{x_1} \cdot \sigma_{x_2} > \left|\sigma_{x_1x_2}\right|$$

$$\left|\sigma_{x_1x_2}\right| < \sigma_{x_1} \cdot \sigma_{x_2} \longrightarrow \begin{cases} \sigma_{x_1x_2} > 0, \left|\sigma_{x_1x_2}\right| = \sigma_{x_1x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \longrightarrow \sigma_{x_1x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \\ \sigma_{x_1x_2} < 0, \left|\sigma_{x_1x_2}\right| = -\sigma_{x_1x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \longrightarrow \sigma_{x_1x_2} > -\sigma_{x_1} \cdot \sigma_{x_2} \end{cases}$$

Finally, the conditions for the matrix Σ (and Σ^{-1}) to be positive definite are:

$$\sigma_{x_1} > 0$$

$$\sigma_{x_2} > 0$$

$$-\sigma_{x_1} \cdot \sigma_{x_2} < \sigma_{x_1x_2} < \sigma_{x_1} \cdot \sigma_{x_2}$$

(Note the lack of the equal symbol)

Reorganizing a bit the terms of the last condition I get:

$$-1 < \frac{\sigma_{x_1 x_2}}{\sigma_{x_1} \cdot \sigma_{x_2}} < 1$$
(Note the lack of the equal symbol)

The correlation coefficient is defined as:

$$\rho = \frac{\sigma_{x_1 x_2}}{\sigma_{x_1} \cdot \sigma_{x_2}}$$
$$-1 \leq \rho \leq 1$$
(Note the presence of the equal symbol)

Basically our conditions stands that I must have strictly positive standard deviations and a correlation coefficient between $(-1, 1)$ (open range, therefore the extremes are excluded):

$$\sigma_{x_1} > 0$$
$$\sigma_{x_2} > 0$$
$$-1 < \rho < 1$$
$$\begin{pmatrix} -\sigma_{x_1} \cdot \sigma_{x_2} < \sigma_{x_1 x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \\ \text{Note the lack of the equal symbol} \end{pmatrix}$$

Let's study the bivariate Gaussian PDF. The covariance matrix Σ is a positive definite matrix with dimension $N \times N$. This means that the covariance matrix Σ is a real symmetric matrix, with dimension $N \times N$, with N real eigenvalues that are strictly positive, $\lambda_j > 0 \; \forall \; j = 1, \ldots, N$, due to the constraints:

$$\begin{gathered} \sigma_{x_1} > 0 \\ \sigma_{x_2} > 0 \\ -1 < \rho < 1 \\ \begin{pmatrix} -\sigma_{x_1} \cdot \sigma_{x_2} < \sigma_{x_1 x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \\ \text{Note the lack of the equal symbol} \end{pmatrix} \end{gathered}$$

and with N orthonormal eigenvectors, $U^T = U^{-1} \longrightarrow U^T \cdot U = I$.

$$\Sigma = \begin{pmatrix} \sigma^2_{x_1} & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma^2_{x_2} \end{pmatrix} = U \cdot D \cdot U^T = \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \sum_{j=1}^{2} \lambda_j \, \vec{u}_j \, \vec{u}_j^T$$

Besides, the matrix Σ^{-1} is also a positive definite matrix with dimension $N \times N$. As before, this means that the matrix Σ^{-1} is a real symmetric matrix, with dimension $N \times N$, with N real eigenvalues that are strictly positive, $1/\lambda_j > 0 \; \forall \; j = 1, \ldots, N$, due to the constraints:

$$\begin{gathered} \sigma_{x_1} > 0 \\ \sigma_{x_2} > 0 \\ -1 < \rho < 1 \\ \begin{pmatrix} -\sigma_{x_1} \cdot \sigma_{x_2} < \sigma_{x_1 x_2} < \sigma_{x_1} \cdot \sigma_{x_2} \\ \text{Note the lack of the equal symbol} \end{pmatrix} \end{gathered}$$

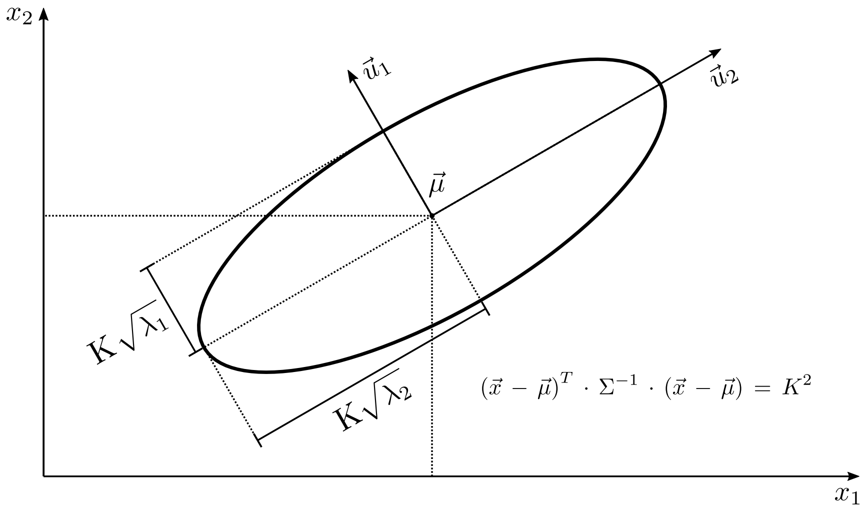
and with N orthonormal eigenvectors, $U^T = U^{-1} \longrightarrow U^T \cdot U = I$.

$$\begin{aligned} \Sigma^{-1} &= \frac{1}{\left|\Sigma\right|} \begin{pmatrix} \sigma^2_{x_2} & -\sigma_{x_1 x_2} \\ -\sigma_{x_1 x_2} & \sigma^2_{x_1} \end{pmatrix} = \frac{1}{\sigma^2_{x_1} \, \sigma^2_{x_2} - \sigma^2_{x_1 x_2}} \begin{pmatrix} \sigma^2_{x_2} & -\sigma_{x_1 x_2} \\ -\sigma_{x_1 x_2} & \sigma^2_{x_1} \end{pmatrix} = \\ &= U \cdot D^{-1} \cdot U^T = \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} 1/\lambda_1 & 0 \\ 0 & 1/\lambda_2 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \\ &= \sum_{j=1}^{2} \frac{1}{\lambda_j} \, \vec{u}_j \, \vec{u}_j^T \end{aligned}$$

The quadratic function

$$(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) = K^2$$

represents a 2D ellipse with its center located at the vector $\vec{\mu}$, the directions of its principal axes defined by the orthonormal eigenvectors of the covariance matrix Σ, $\vec{u}_1$ and $\vec{u}_2$, and with the length of its principal axes defined by $2\,K\,\sqrt{\lambda_j}$.



The quadratic function $(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu})$ can be express in two ways:

$$
\begin{aligned}
(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) &= \frac{1}{\left|\Sigma\right|} \cdot \Big(x_1 - \mu_1, x_2 - \mu_2\Big) \cdot \begin{pmatrix} \sigma^2_{x_2} & -\sigma_{x_1 x_2} \\ -\sigma_{x_1 x_2} & \sigma^2_{x_1} \end{pmatrix} \cdot \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} = \\
&= \frac{1}{\left|\Sigma\right|} \cdot \Big(\sigma^2_{x_2}\,(x_1 - \mu_1)^2 + \sigma^2_{x_1}\,(x_2 - \mu_2)^2 - 2\,\sigma_{x_1 x_2}\,(x_1 - \mu_1)\,(x_2 - \mu_2)\Big) \\
(\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) &= \sum_{j=1}^{2} \frac{1}{\lambda_j} \cdot \left(\vec{u}_j^{\,T} \cdot (\vec{x} - \vec{\mu})\right)^2 = \\
&= \left(\frac{u_{11}^2}{\lambda_1} + \frac{u_{21}^2}{\lambda_2}\right) \cdot (x_1 - \mu_1)^2 + \left(\frac{u_{12}^2}{\lambda_1} + \frac{u_{22}^2}{\lambda_2}\right) \cdot (x_2 - \mu_2)^2 + \\
&+ 2 \cdot \left(\frac{u_{11} \cdot u_{12}}{\lambda_1} + \frac{u_{21} \cdot u_{22}}{\lambda_2}\right) \cdot (x_1 - \mu_1) \cdot (x_2 - \mu_2)
\end{aligned}
$$

$$
\begin{aligned}
Pr\,(\vec{x} \in \text{ ellipse of } K = 1) &= 0.682689 \;\; (68.269\%) \\
Pr\,(\vec{x} \in \text{ ellipse of } K = 2) &= 0.954499 \;\; (95.450\%) \\
Pr\,(\vec{x} \in \text{ ellipse of } K = 3) &= 0.997300 \;\; (99.730\%)
\end{aligned}
$$

Let's describe how to calculate the eigenvalues and the eigenvectors of the matrix Σ.
If the following expression can be written:

$$\Sigma \cdot \vec{u}_j = \lambda_j \cdot \vec{u}_j$$

then value λ_j is an eigenvalue of the matrix Σ and the vector $\vec{u}_j$ is an eigenvector of the matrix Σ that are

associated.
Therefore, if the above equation is satisfied, then the equation can be written as:

$$\Sigma \cdot \vec{u}_j - \lambda_j \cdot \vec{u}_j = (\Sigma - \lambda_j I) \cdot \vec{u}_j = 0$$

The equation $(\Sigma - \lambda_j I) \cdot \vec{u}_j = 0$ has a non-zero solution, $\vec{u}_j \neq 0$, if and only if $\left|\Sigma - \lambda_j I\right| = 0$. Therefore, the eigenvalues of the covariance matrix Σ are the roots of the equation $\left|\Sigma - \lambda_j I\right| = 0$. The values λ_j obtained are introduced in the expression $(\Sigma - \lambda_j I) \cdot \vec{u}_j = 0$ to calculate the eigenvectors $\vec{u}_j$.
We are going to use the relation:

$$\rho = \frac{\sigma_{x_1 x_2}}{\sigma_{x_1}\, \sigma_{x_2}}$$
$$-1 < \rho < 1$$

$\Big($Note the lack of the equal symbol due to our conditions for the matrices Σ and Σ^{-1} to be posititive definite$\Big)$

$$\begin{aligned}
\left|\Sigma - \lambda_j I\right| &= \left|\begin{pmatrix} \sigma^2_{x_1} & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma^2_{x_2} \end{pmatrix} - \begin{pmatrix} \lambda_j & 0 \\ 0 & \lambda_j \end{pmatrix}\right| = \left|\begin{pmatrix} \sigma^2_{x_1} - \lambda_j & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma^2_{x_2} - \lambda_j \end{pmatrix}\right| = \\
&= \left(\sigma^2_{x_1} - \lambda_j\right)\left(\sigma^2_{x_2} - \lambda_j\right) - \sigma^2_{x_1 x_2} = \\
&= \left(\sigma^2_{x_1} - \lambda_j\right)\left(\sigma^2_{x_2} - \lambda_j\right) - \rho^2 \sigma^2_{x_1} \sigma^2_{x_2} = \\
&= \lambda_j^2 - \left(\sigma^2_{x_1} + \sigma^2_{x_2}\right)\lambda_j + \sigma^2_{x_1}\, \sigma^2_{x_2}\left(1 - \rho^2\right) = 0
\end{aligned}$$

$$\begin{aligned}
\lambda_1 &= \frac{1}{2}\left(\sigma^2_{x_1} + \sigma^2_{x_2} + \sqrt{\left(\sigma^2_{x_1} + \sigma^2_{x_2}\right)^2 - 4\,\sigma^2_{x_1}\,\sigma^2_{x_2}\,(1 - \rho^2)}\right) \\
\lambda_2 &= \frac{1}{2}\left(\sigma^2_{x_1} + \sigma^2_{x_2} - \sqrt{\left(\sigma^2_{x_1} + \sigma^2_{x_2}\right)^2 - 4\,\sigma^2_{x_1}\,\sigma^2_{x_2}\,(1 - \rho^2)}\right)
\end{aligned}$$

Now, I calculate the eigenvectors $\vec{u}_j$:

$$\Sigma \cdot \vec{u}_1 = \lambda_1 \vec{u}_1$$

$$\begin{pmatrix} \sigma^2_{x_1} & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma^2_{x_2} \end{pmatrix} \cdot \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \lambda_1 \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix}$$

$$\begin{pmatrix} \sigma^2_{x_1} & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma^2_{x_2} \end{pmatrix} \cdot \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} - \lambda_1 \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = 0$$

$$\left(\sigma_{x_1}^2 - \lambda_1\right) u_{11} + \sigma_{x_1 x_2} u_{12} = 0$$

$$\sigma_{x_1 x_2} u_{11} + \left(\sigma_{x_2}^2 - \lambda_1\right) u_{12} = 0$$

$$\rho = \frac{\sigma_{x_1 x_2}}{\sigma_{x_1} \sigma_{x_2}}$$

$$\left(\sigma_{x_1}^2 - \lambda_1\right) u_{11} + \rho \sigma_{x_1} \sigma_{x_2} u_{12} = 0$$

$$\rho \sigma_{x_1} \sigma_{x_2} u_{11} + \left(\sigma_{x_2}^2 - \lambda_1\right) u_{12} = 0$$

$$\begin{aligned} \sigma_{x_1}^2 - \lambda_1 &= \sigma_{x_1}^2 - \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4 \sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}\right) \\ &= \frac{1}{2}\left(\sigma_{x_1}^2 - \sigma_{x_2}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4 \sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}\right) \end{aligned}$$

$$\begin{aligned} u_{12} &= \frac{-\left(\sigma_{x_1}^2 - \lambda_1\right)}{\rho \sigma_{x_1} \sigma_{x_2}} u_{11} = \\ &= \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4 \sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}}{2 \rho \sigma_{x_1} \sigma_{x_2}} u_{11} \end{aligned}$$

Solution for the vector $\vec{u}_1$:

$$\vec{u}_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \begin{pmatrix} d \\ d\, b \end{pmatrix}$$

where the term d is any scalar value and b is:

$$b = \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4 \sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}}{2 \rho \sigma_{x_1} \sigma_{x_2}}$$

Choose d to satisfy:

$$\|\vec{u}_1\|^2 = \vec{u}_1^T \cdot \vec{u}_1 = 1$$

$$u_{11}^2 + u_{12}^2 = d^2 + d^2 b^2 = 1$$
$$d^2 \left(1 + b^2\right) = 1$$
$$d^2 = \frac{1}{(1 + b^2)}$$
$$\sqrt{d^2} = \frac{1}{\sqrt{(1 + b^2)}}$$
$$\left|d\right| = \frac{1}{\sqrt{(1 + b^2)}}$$
$$d > 0, \left|d\right| = d = \frac{1}{\sqrt{(1 + b^2)}} \qquad \longrightarrow d = \frac{1}{\sqrt{(1 + b^2)}}$$
$$d < 0, \left|d\right| = -d = \frac{1}{\sqrt{(1 + b^2)}} \qquad \longrightarrow d = \frac{-1}{\sqrt{(1 + b^2)}}$$

So finally I have the orthonormal eigenvector $\vec{u}_1$:

$$\vec{u}_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \begin{pmatrix} d \\ d\,b \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+b^2}} \\ \frac{b}{\sqrt{1+b^2}} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{-1}{\sqrt{1+b^2}} \\ \frac{-b}{\sqrt{1+b^2}} \end{pmatrix}$$

Now, I repeat the same procedure to get the orthonormal eigenvector $\vec{u}_2$:

$$\Sigma \cdot \vec{u}_2 = \lambda_2 \vec{u}_2$$

$$\begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma_{x_2}^2 \end{pmatrix} \cdot \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \lambda_2 \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix}$$

$$\begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1 x_2} \\ \sigma_{x_1 x_2} & \sigma_{x_2}^2 \end{pmatrix} \cdot \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} - \lambda_2 \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = 0$$

$$\left(\sigma_{x_1}^2 - \lambda_2\right) u_{21} + \sigma_{x_1 x_2} u_{22} = 0$$

$$\sigma_{x_1 x_2} u_{21} + \left(\sigma_{x_2}^2 - \lambda_2\right) u_{22} = 0$$

$$\rho = \frac{\sigma_{x_1 x_2}}{\sigma_{x_1} \sigma_{x_2}}$$

$$\left(\sigma_{x_1}^2 - \lambda_2\right) u_{21} + \rho \, \sigma_{x_1} \sigma_{x_2} u_{22} = 0$$

$$\rho \, \sigma_{x_1} \sigma_{x_2} u_{21} + \left(\sigma_{x_2}^2 - \lambda_2\right) u_{22} = 0$$

$$\begin{aligned} \sigma_{x_1}^2 - \lambda_2 &= \sigma_{x_1}^2 - \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}\right) \\ &= \frac{1}{2}\left(\sigma_{x_1}^2 - \sigma_{x_2}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}\right) \end{aligned}$$

$$\begin{aligned} u_{22} &= \frac{-\left(\sigma_{x_1}^2 - \lambda_2\right)}{\rho \, \sigma_{x_1} \sigma_{x_2}} u_{21} = \\ &= \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}}{2\,\rho \, \sigma_{x_1} \sigma_{x_2}} u_{21} \end{aligned}$$

Solution for the vector $\vec{u}_2$:

$$\vec{u}_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \begin{pmatrix} e \\ e\,c \end{pmatrix}$$

where the term e is any scalar value and c is:

$$c = \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2 \sigma_{x_2}^2 \left(1 - \rho^2\right)}}{2\,\rho \, \sigma_{x_1} \sigma_{x_2}}$$

Choose e to satisfy:

$$\|\vec{u}_2\|^2 = \vec{u}_2^T \cdot \vec{u}_2 = 1$$

$$
\begin{aligned}
u_{21}^2 + u_{22}^2 &= e^2 + e^2\,c^2 = 1 \\
e^2\left(1 + c^2\right) &= 1 \\
e^2 &= \frac{1}{(1 + c^2)} \\
\sqrt{e^2} &= \frac{1}{\sqrt{(1 + c^2)}} \\
\left|e\right| &= \frac{1}{\sqrt{(1 + c^2)}} \\
e > 0,\ \left|e\right| = e = \frac{1}{\sqrt{(1 + c^2)}} \quad &\longrightarrow e = \frac{1}{\sqrt{(1 + c^2)}} \\
e < 0,\ \left|e\right| = -\,e = \frac{1}{\sqrt{(1 + c^2)}} \quad &\longrightarrow e = \frac{-1}{\sqrt{(1 + c^2)}}
\end{aligned}
$$

So finally I have:

$$
\vec{u}_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \begin{pmatrix} e \\ e\,c \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+c^2}} \\ \frac{c}{\sqrt{1+b^2}} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{-\,1}{\sqrt{1+c^2}} \\ \frac{-\,c}{\sqrt{1+c^2}} \end{pmatrix}
$$

Summary:

$$
\begin{gathered}
\sigma_{x_1} > 0 \\
\sigma_{x_2} > 0 \\
-\,1 < \rho < 1 \\
\Big(\text{Note the lack of the equal symbol}\Big) \\
\sigma_{x_1 x_2} = \rho\,\sigma_{x_1}\,\sigma_{x_2}
\end{gathered}
$$

$$
\begin{aligned}
\lambda_1 &= \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2\,\sigma_{x_2}^2\,(1 - \rho^2)}\right) \\
\lambda_2 &= \frac{1}{2}\left(\sigma_{x_1}^2 + \sigma_{x_2}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2\,\sigma_{x_2}^2\,(1 - \rho^2)}\right)
\end{aligned}
$$

$$
b = \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 + \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2\,\sigma_{x_2}^2\,(1 - \rho^2)}}{2\,\rho\,\sigma_{x_1}\,\sigma_{x_2}}
$$

$$
c = \frac{\sigma_{x_2}^2 - \sigma_{x_1}^2 - \sqrt{\left(\sigma_{x_1}^2 + \sigma_{x_2}^2\right)^2 - 4\,\sigma_{x_1}^2\,\sigma_{x_2}^2\,(1 - \rho^2)}}{2\,\rho\,\sigma_{x_1}\,\sigma_{x_2}}
$$

$$\vec{u}_1 = \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+b^2}} \\ \frac{b}{\sqrt{1+b^2}} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{-1}{\sqrt{1+b^2}} \\ \frac{-b}{\sqrt{1+b^2}} \end{pmatrix}$$

$$\vec{u}_2 = \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+c^2}} \\ \frac{c}{\sqrt{1+c^2}} \end{pmatrix} \text{ or } \begin{pmatrix} \frac{-1}{\sqrt{1+c^2}} \\ \frac{-c}{\sqrt{1+c^2}} \end{pmatrix}$$

The vector $\vec{u}_1$ and the vector $\vec{u}_2$ are orthonormal:

$$U^T \cdot U = I \longrightarrow U^T = U^{-1}$$

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

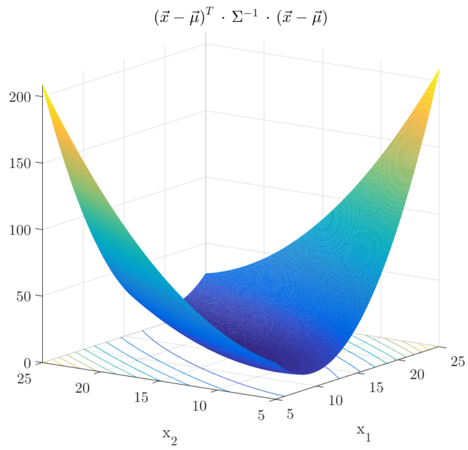
$$\|\vec{u}_1\|^2 = \langle \vec{u}_1, \vec{u}_1 \rangle = \vec{u}_1^T \cdot \vec{u}_1 = u_{11}^2 + u_{12}^2 = 1$$

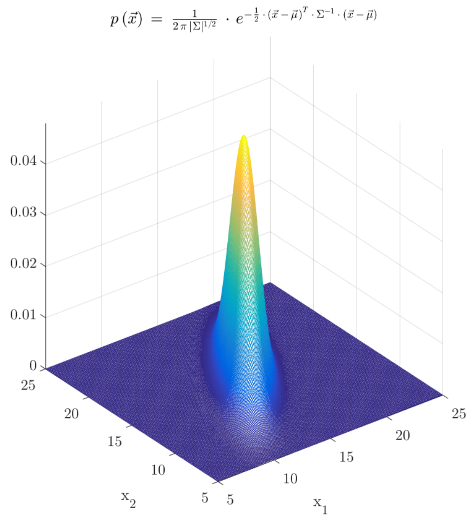
$$\|\vec{u}_2\|^2 = \langle \vec{u}_2, \vec{u}_2 \rangle = \vec{u}_2^T \cdot \vec{u}_2 = u_{21}^2 + u_{22}^2 = 1$$

$$\langle \vec{u}_1, \vec{u}_2 \rangle = \vec{u}_1^T \cdot \vec{u}_2 = u_{11}\, u_{21} + u_{12}\, u_{22} = 0$$

Examples:

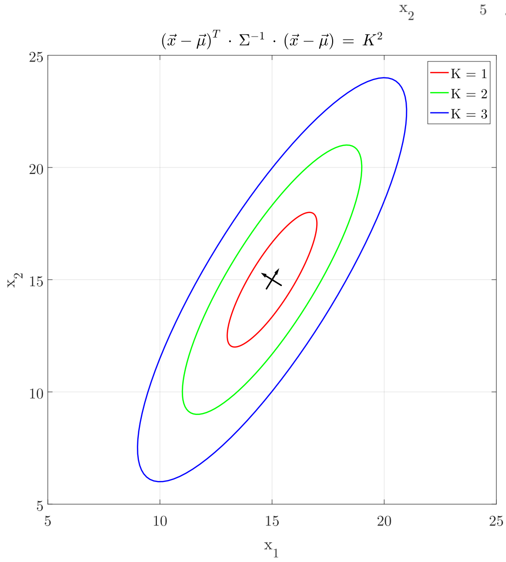
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 2^2 & 5 \\ 5 & 3^2 \end{pmatrix}$$



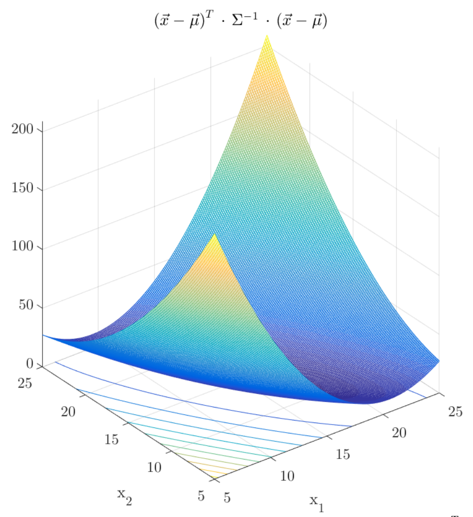


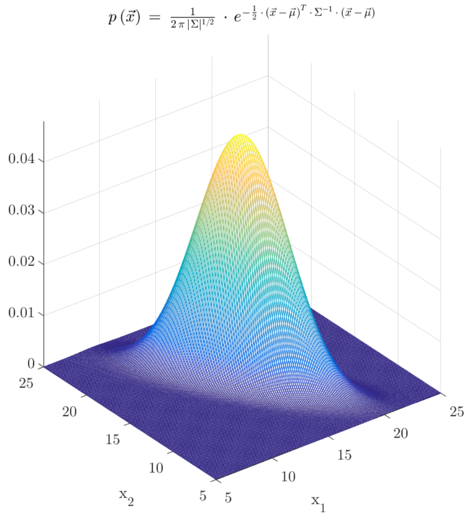
$$U = \begin{pmatrix} -0.851 & +0.526 \\ +0.526 & +0.851 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.910 & 0 \\ 0 & 12.090 \end{pmatrix}$$



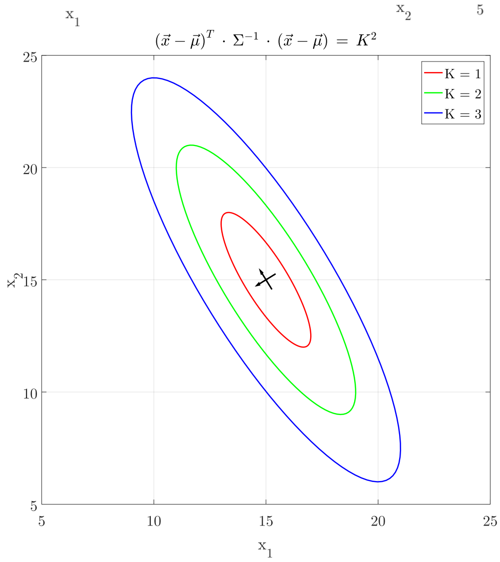
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 2^2 & -5 \\ -5 & 3^2 \end{pmatrix}$$



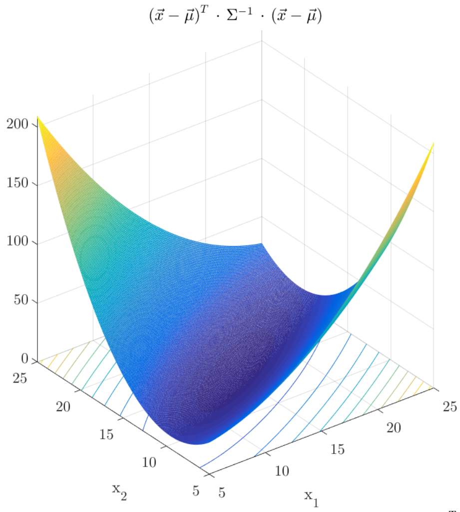


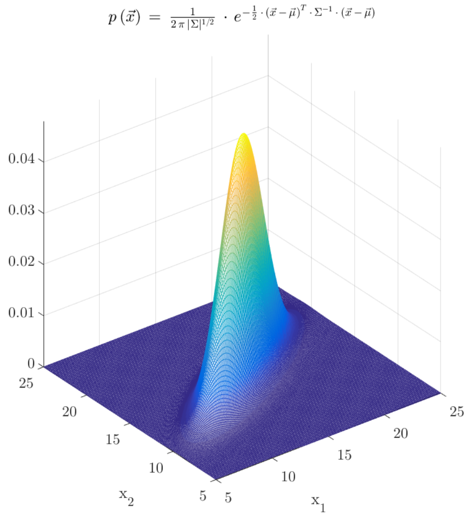
$$U = \begin{pmatrix} -0.851 & -0.526 \\ -0.526 & +0.851 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.910 & 0 \\ 0 & 12.090 \end{pmatrix}$$



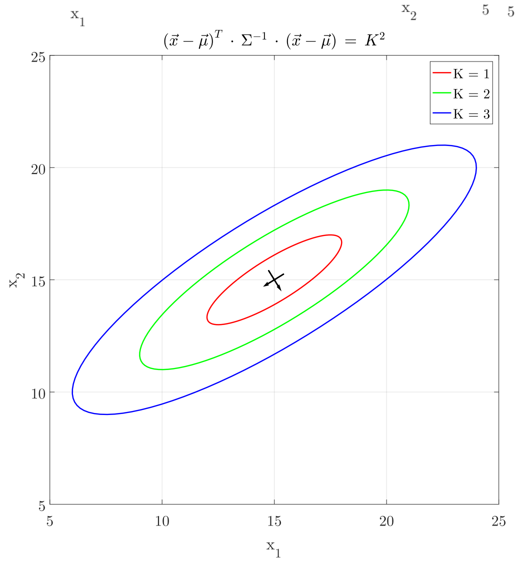
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \; \Sigma = \begin{pmatrix} 3^2 & 5 \\ 5 & 2^2 \end{pmatrix}$$



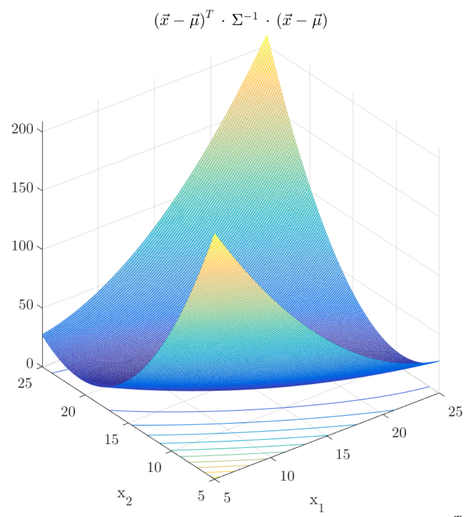


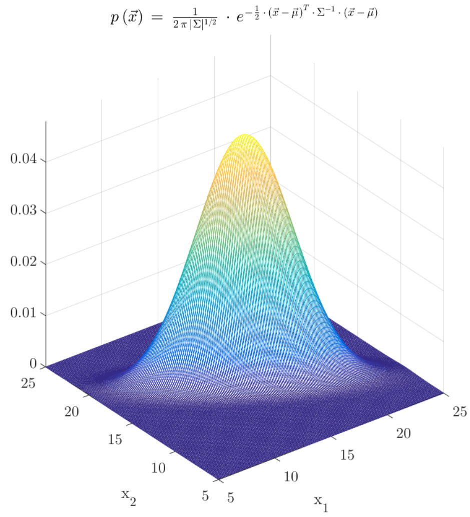
$$U = \begin{pmatrix} +0.526 & -0.851 \\ -0.851 & -0.526 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.910 & 0 \\ 0 & 12.090 \end{pmatrix}$$



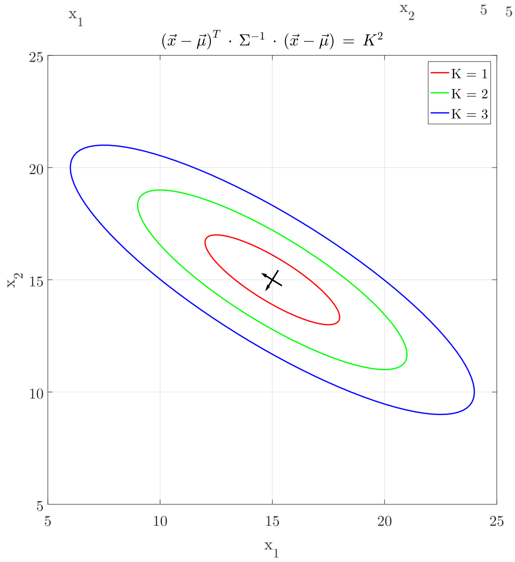
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 3^2 & -5 \\ -5 & 2^2 \end{pmatrix}$$





$$U = \begin{pmatrix} -0.526 & -0.851 \\ -0.851 & +0.526 \end{pmatrix}$$

$$D = \begin{pmatrix} 0.910 & 0 \\ 0 & 12.090 \end{pmatrix}$$

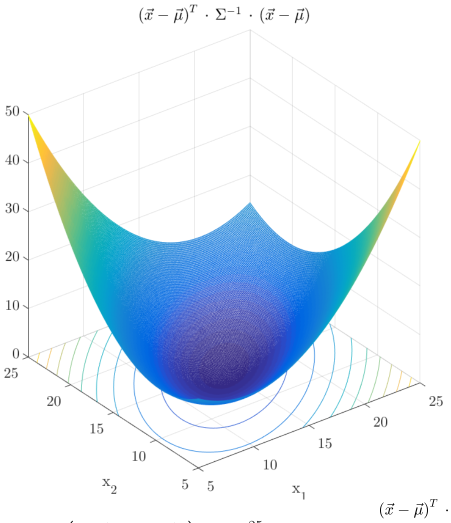


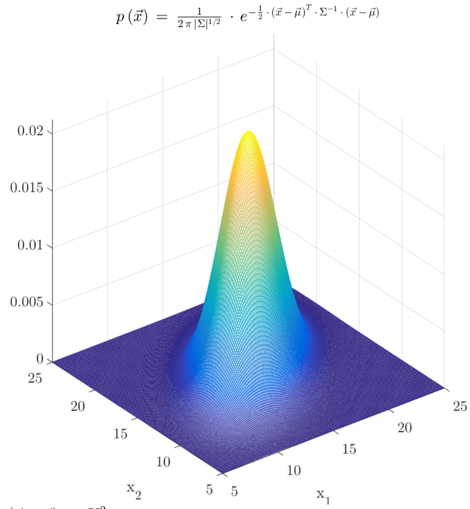
If $\sigma_{x_1} = \sigma_{x_2}$

$$\begin{aligned}
\vec{u}_1 &= \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ +\frac{1}{\sqrt{2}} \end{pmatrix} \text{ or } \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \\
\vec{u}_2 &= \begin{pmatrix} +\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \text{ or } \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ +\frac{1}{\sqrt{2}} \end{pmatrix} \\
\lambda_1 &= \sigma_{x_1}^2 (1 + \rho) = \sigma_{x_1}^2 + \sigma_{x_1 x_2} \\
\lambda_2 &= \sigma_{x_1}^2 (1 - \rho) = \sigma_{x_1}^2 - \sigma_{x_1 x_2}
\end{aligned}$$

Therefore, if $\sigma_{x_1 x_2} > 0$, then the major axis of the ellipse will be in the direction of the 45° line and the minor axis of the ellipse will be in the direction of the 135° (45° + 90°) line.
Therefore, if $\sigma_{x_1 x_2} < 0$, then the major axis of the ellipse will be in the direction of the 135° line and the minor axis of the ellipse will be in the direction of the 45° (135° − 90°) line.

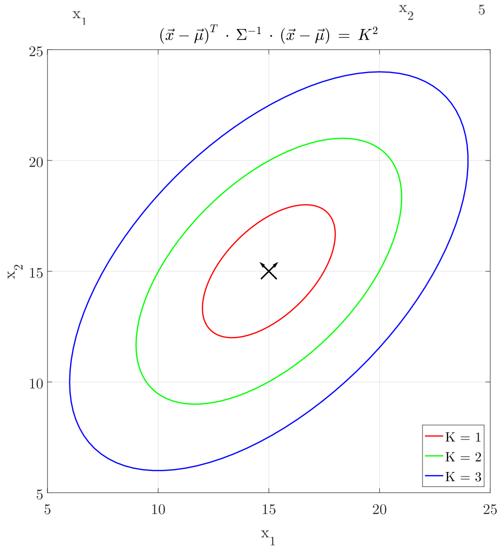
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \; \Sigma = \begin{pmatrix} 3^2 & 5 \\ 5 & 3^2 \end{pmatrix}$$



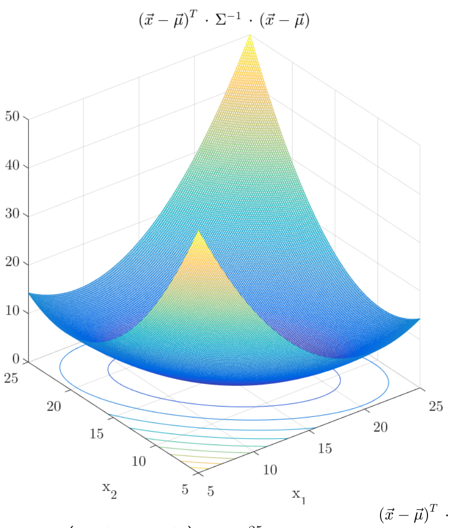


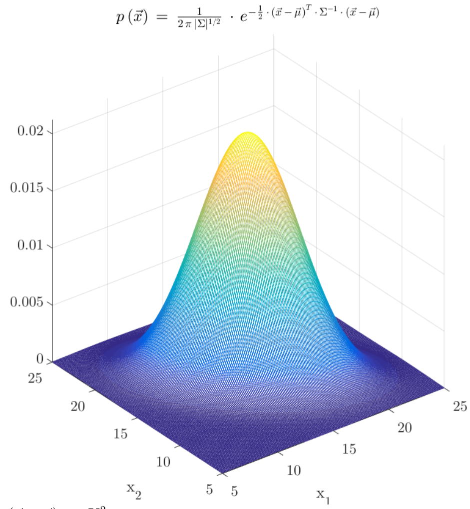
$$U = \begin{pmatrix} +\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ +\frac{1}{\sqrt{2}} & +\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$D = \begin{pmatrix} 14 & 0 \\ 0 & 4 \end{pmatrix}$$



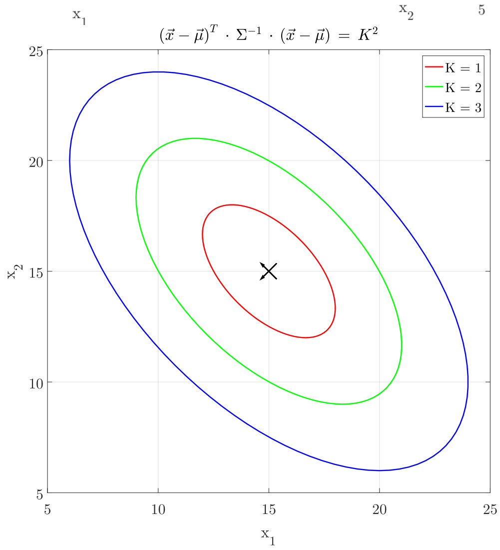
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \; \Sigma = \begin{pmatrix} 3^2 & -5 \\ -5 & 3^2 \end{pmatrix}$$





$$U = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & +\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$D = \begin{pmatrix} 4 & 0 \\ 0 & 14 \end{pmatrix}$$



Let's describe briefly how to plot a 2D ellipse given the directions and lengths of its two principal axes. The directions of the ellipse's two principal axes are given by the orthonormal eigenvectors $\vec{u}_1$ and $\vec{u}_2$. The length of the ellipse's two principal axes are given by $2\,K\,\sqrt{\lambda_1}$ and $2\,K\,\sqrt{\lambda_2}$.

$$U = \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$\begin{aligned}
\begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_j & y_j \\ \vdots & \vdots \\ x_M & y_M \end{pmatrix} &= K \begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \vdots & \vdots \\ \cos(\theta_j) & \sin(\theta_j) \\ \vdots & \vdots \\ \cos(\theta_M) & \sin(\theta_M) \end{pmatrix} \cdot \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \\
&= K \begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \vdots & \vdots \\ \cos(\theta_j) & \sin(\theta_j) \\ \vdots & \vdots \\ \cos(\theta_M) & \sin(\theta_M) \end{pmatrix} \cdot \begin{pmatrix} \sqrt{\lambda_1}\,u_{11} & \sqrt{\lambda_1}\,u_{12} \\ \sqrt{\lambda_2}\,u_{21} & \sqrt{\lambda_2}\,u_{22} \end{pmatrix} = \\
&= K \begin{pmatrix} \sqrt{\lambda_1}\,u_{11}\,\cos(\theta_1) + \sqrt{\lambda_2}\,u_{21}\,\sin(\theta_1) & \sqrt{\lambda_1}\,u_{12}\,\cos(\theta_1) + \sqrt{\lambda_2}\,u_{22}\,\sin(\theta_1) \\ \vdots & \vdots \\ \sqrt{\lambda_1}\,u_{11}\,\cos(\theta_j) + \sqrt{\lambda_2}\,u_{21}\,\sin(\theta_j) & \sqrt{\lambda_1}\,u_{12}\,\cos(\theta_j) + \sqrt{\lambda_2}\,u_{22}\,\sin(\theta_j) \\ \vdots & \vdots \\ \sqrt{\lambda_1}\,u_{11}\,\cos(\theta_M) + \sqrt{\lambda_2}\,u_{21}\,\sin(\theta_M) & \sqrt{\lambda_1}\,u_{12}\,\cos(\theta_M) + \sqrt{\lambda_2}\,u_{22}\,\sin(\theta_M) \end{pmatrix}
\end{aligned}$$

$$\theta_j = \frac{360}{M}\,j$$
$$j = 0, \ldots, M - 1$$
$$\Big(\text{M is the desired number of points to draw}\Big)$$

Now, I'm going to describe one particular case that occurs when $\sigma_{x_1 x_2} = 0$, which it's the same as saying $\rho = 0$.
When

$$\sigma_{x_1 x_2} = \rho\,\sigma_{x_1}\,\sigma_{x_2} \neq 0$$
$$\Big(-1 < \rho < 1\Big)$$

the variables x_1 and x_2 are correlated.
When

$$\sigma_{x_1 x_2} = \rho \, \sigma_{x_1} \sigma_{x_2} = 0$$
$$\Big(\rho = 0 \Big)$$

the variables x_1 and x_2 are incorrelated. Now, one interesting result stands that if the joint distribution among N random variables is Gaussian, then the marginal distribution of each random variable is also Gaussian, even the conditional distributions are Gaussian. The other way around is not true. If you have N random variables, each of those distributed following a Gaussian PDF, then the random vector doesn't have to be distributed following a Gaussian PDF [1]
So because the joint distribution between the random variables x_1 and x_2 is Gaussian, then the random variables x_1 and x_2 are also distributed accordingly to Gaussian PDFs. For random variables distributed accordingly to Gaussian distributions incorrelation implies independency.
Every computation is easier when $\sigma_{x_1 x_2} = 0$.

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & 0 & \dots & 0 & \dots & 0 \\ 0 & \sigma_{x_2}^2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_{x_j}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & \sigma_{x_N}^2 \end{pmatrix} = \sum_{j=1}^{N} \lambda_j \, \vec{e}_j \, \vec{e}_j^{\,T}$$

$$\vec{e}_j = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longleftarrow j - th \text{ position}$$

$$\lambda_j = \sigma_{x_j}^2$$

$$\Sigma^{-1} = \sum_{j=1}^{N} \frac{1}{\lambda_j} \, \vec{e}_j \, \vec{e}_j^{\,T} = \begin{pmatrix} 1/\sigma_{x_1}^2 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1/\sigma_{x_2}^2 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1/\sigma_{x_j}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1/\sigma_{x_N}^2 \end{pmatrix}$$

[1] Not 100% sure about this last fact. If you ever need this result, check it out firstly.

Particularizing for the 2D Gaussian:

$$\Sigma = \begin{pmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{pmatrix}$$

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_{x_1}^2} & 0 \\ 0 & \frac{1}{\sigma_{x_2}^2} \end{pmatrix}$$

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\lambda_1 = \sigma_{x_1}^2 \quad \lambda_2 = \sigma_{x_2}^2$$

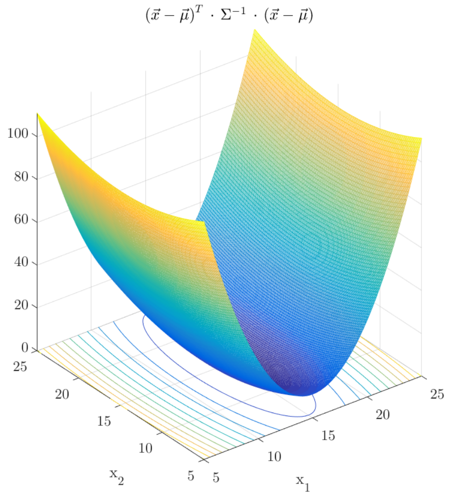
$$\begin{aligned}
\left\langle (\vec{x} - \vec{\mu}),\, \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right\rangle &= (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) = \\
&= \begin{pmatrix} x_1 - \mu_1, x_2 - \mu_2 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sigma_{x_1}^2} & 0 \\ 0 & \frac{1}{\sigma_{x_2}^2} \end{pmatrix} \cdot \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} = \\
&= \frac{(x_1 - \mu_1)^2}{\sigma_{x_1}^2} + \frac{(x_2 - \mu_2)^2}{\sigma_{x_2}^2} \\
&= \frac{e_{11}^2}{\lambda_1} \cdot (x_1 - \mu_1)^2 + \frac{e_{22}^2}{\lambda_2} \cdot (x_2 - \mu_2)^2
\end{aligned}$$

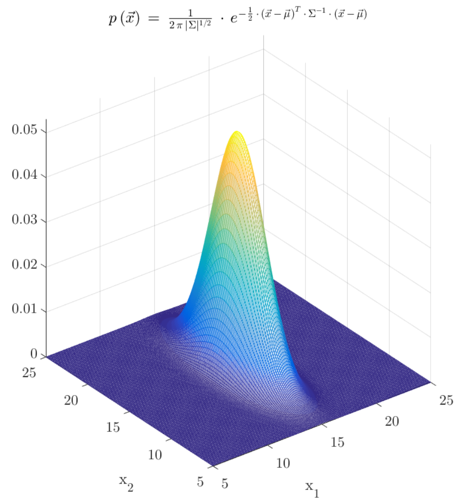
The following expression corresponds to an ellipse centered at the vector $\vec{\mu}$, with the directions of its principal axes defined by the orthonormal eigenvectors $\vec{e}_1$ and $\vec{e}_2$, i.e, they are parallels to the x-axis and y-axis, and with the length of its principal axes defined by $K\sigma_{x_1}$.

$$\begin{aligned}
\left\langle (\vec{x} - \vec{\mu}),\, \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) \right\rangle &= (\vec{x} - \vec{\mu})^T \cdot \Sigma^{-1} \cdot (\vec{x} - \vec{\mu}) = \\
&= \frac{(x_1 - \mu_1)^2}{\sigma_{x_1}^2} + \frac{(x_2 - \mu_2)^2}{\sigma_{x_2}^2} = K^2
\end{aligned}$$

$$\left(\frac{x_1 - \mu_1}{K\,\sigma_{x_1}} \right)^2 + \left(\frac{x_2 - \mu_2}{K\,\sigma_{x_2}} \right)^2 = 1$$

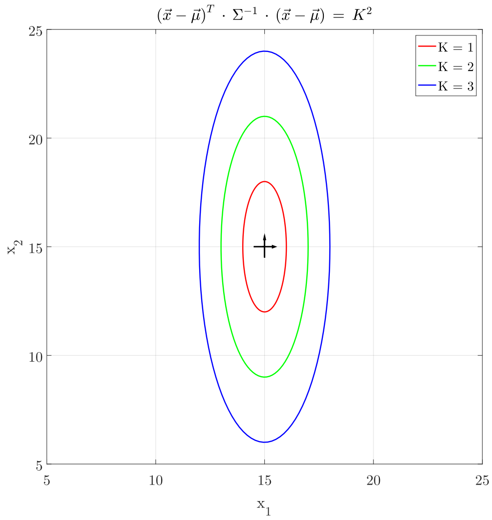
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 1^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$



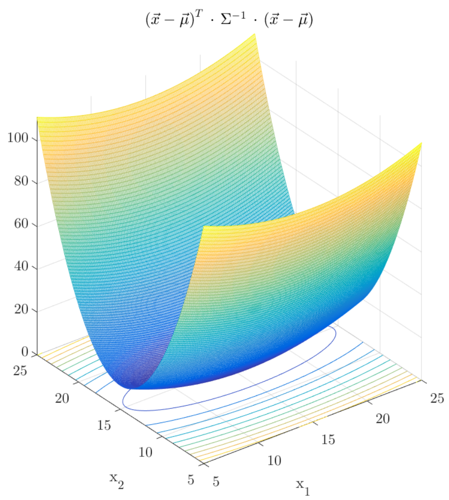


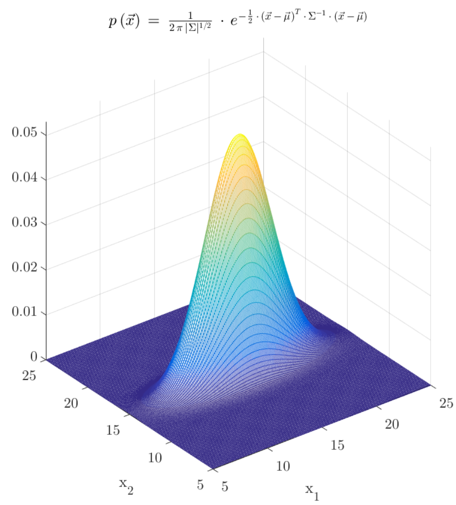
$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} 1^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$



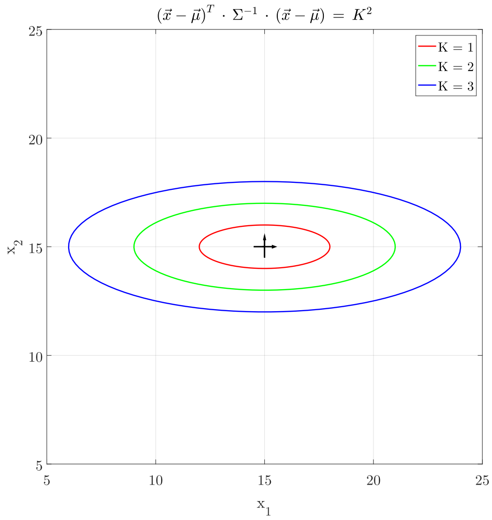
$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \ \Sigma = \begin{pmatrix} 3^2 & 0 \\ 0 & 1^2 \end{pmatrix}$$



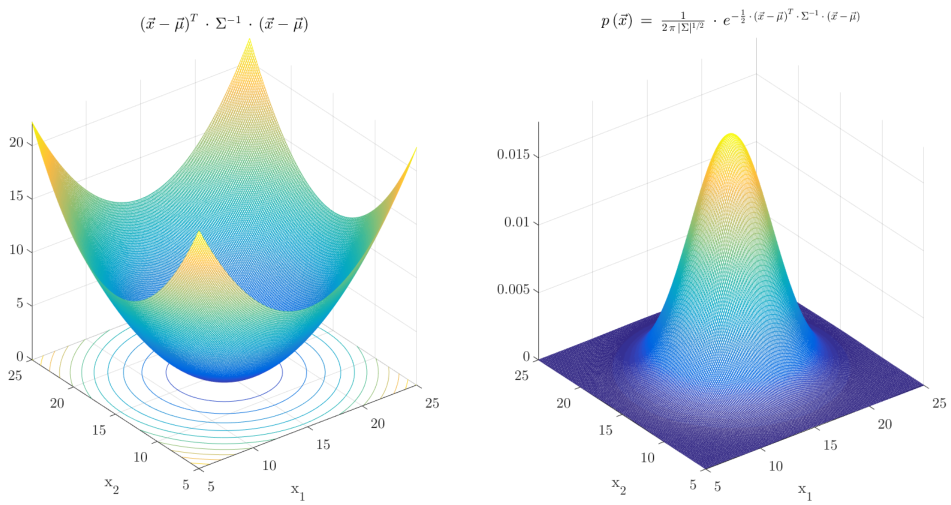


$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

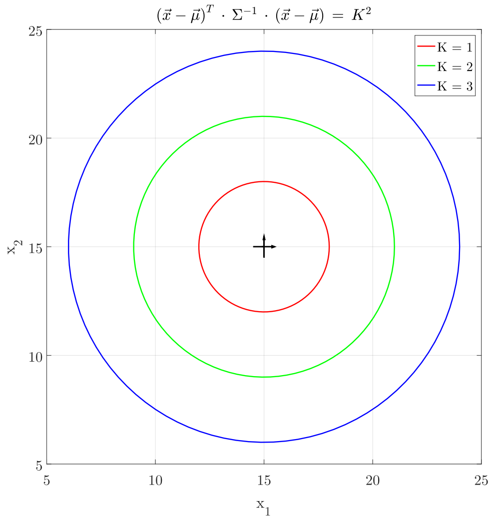
$$D = \begin{pmatrix} 3^2 & 0 \\ 0 & 1^2 \end{pmatrix}$$



$$\mu = \begin{pmatrix} 15 \\ 15 \end{pmatrix} \ \Sigma = \begin{pmatrix} 3^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$



$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$D = \begin{pmatrix} 3^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$



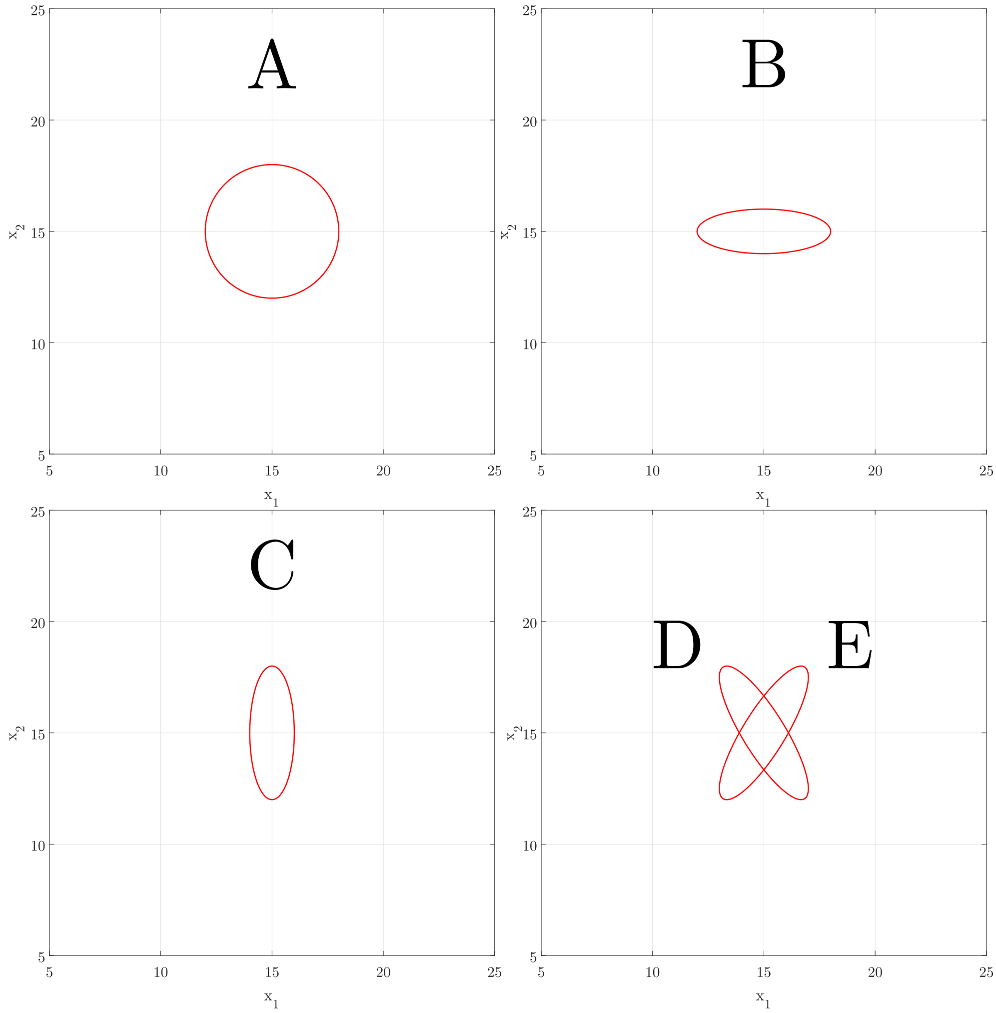
So, as we can see, when $\sigma_{x_1 x_2} = 0$:

- If $\sigma_{x_1} > \sigma_{x_2}$ then the ellipse has its major axis along the x-axis.
- If $\sigma_{x_2} > \sigma_{x_1}$ then the ellipse has its major axis along the y-axis.
- If $\sigma_{x_1} = \sigma_{x_2}$ then the ellipse converts into a circumference, and its radius is equal $r = K\sigma_{x_1} = K\sigma_{x_2}$.

Exercise:
Let's consider a brief exercise. Let's suppose that we are partially given a Gaussian joint distribution between the random variable x_1 and the random variable x_2.

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad p(\vec{x}) = \mathcal{N}(\vec{\mu}, \Sigma) \qquad \vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma^2 & ? \\ ? & \sigma^2 \end{pmatrix}$$

With the given partial Gaussian joint distribution let's try to associate it with one of the pictures below.



First thing $\sigma^2_{x_1} = \sigma^2_{x_2} = \sigma^2$.

If $\sigma_{x_1 x_2} = 0$, then we would have a circumference. So, figure A.

If $\sigma_{x_1 x_2} \neq 0$, then we would have an ellipse. If $\sigma_{x_1 x_2} > 0$, then the major axis of the ellipse will be in the direction of the 45° line and the minor axis of the ellipse will be in the direction of the 135° (45° + 90°) line. So, figure E.

If $\sigma_{x_1 x_2} < 0$, then the major axis of the ellipse will be in the direction of the 135° line and the minor axis of the ellipse will be in the direction of the 45° (135° − 90°) line. So, figure D.

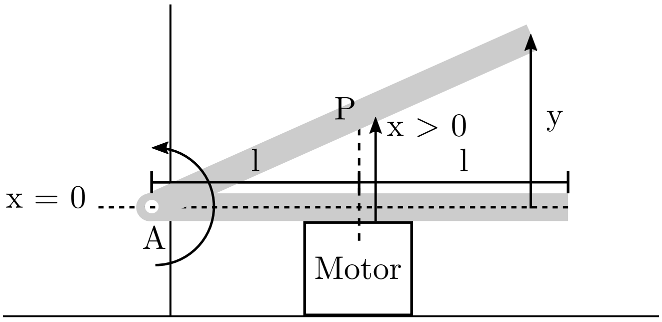
We would have the figure B when $\sigma^2_{x_1} > \sigma^2_{x_2}$ and $\sigma_{x_1 x_2} = 0$.

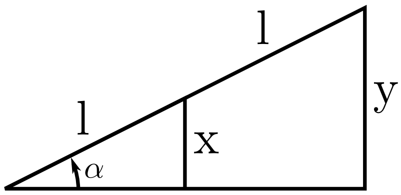
We would have the figure C when $\sigma^2_{x_1} < \sigma^2_{x_2}$ and $\sigma_{x_1 x_2} = 0$.

Therefore, since we have $\sigma^2_{x_1} = \sigma^2_{x_2} = \sigma^2$ we can't have either figure B or figure C.

EXERCISE

The lever arm is fixed to the wall at the point A. The lever arm can rotate around the point A. The motor moves the point P a distance x. First, calculate the relation between the random variable y and the random variable x. Calculate the eigenvalues and the eigenvectors of the covariance matrix Σ that results from the relation between the random variables x and y.





The free extreme of the lever arm moves a distance:
gather

$$\sin(\alpha) = \frac{x}{l}$$
$$y = 2l \sin(\alpha) = 2l\,\frac{x}{l} = 2x$$

We are told that the error in the motor motion, i.e, the error in the x distance is:

$$e_x = \mathcal{N}\left(\mu, \sigma_x^2\right)$$
$$\mu = 0$$

The distance y only depends on the motor, so, because the motor has an inherent error e_x that is distributed according to a Gaussian distribution, the error in the motion in the free extreme of the lever arm also follows a Gaussian distribution. Let's suppose that the motor moves a distance x with some error:

$$
\begin{aligned}
x' &= x + e_x \\
y' &= y + e_y = 2x' = 2\left(x + e_x\right) = 2x + 2e_x \\
y &= 2x \\
e_y &= 2e_x
\end{aligned}
$$

$$
\begin{aligned}
\mu_x &\triangleq E\left(e_x\right) = 0 \\
\sigma_x^2 &\triangleq E\left(\left(e_x - E\left(e_x\right)\right)^2\right) = E\left(e_x^2\right) \\
\mu_y &\triangleq E\left(e_y\right) = E\left(2e_x\right) = 2E\left(e_x\right) = 0 \\
\sigma_y^2 &\triangleq E\left(\left(e_y - E\left(e_y\right)\right)^2\right) = E\left(e_y^2\right) = E\left(4e_x^2\right) = 4E\left(e_x^2\right) = 4\sigma_x^2
\end{aligned}
$$

$$
e_y = \mathcal{N}\left(0,\, 4\sigma_x^2\right)
$$

$$
\vec{e} = \begin{pmatrix} e_x \\ e_y \end{pmatrix}
$$

$$
\begin{aligned}
\Sigma &= E\left(\left(\vec{e} - E\left(\vec{e}\right)\right)\left(\vec{e} - E\left(\vec{e}\right)\right)^T\right) = E\left(\vec{e} \cdot \vec{e}^T\right) = E\left(\begin{pmatrix} e_x \\ e_y \end{pmatrix} \cdot \begin{pmatrix} e_x & e_y \end{pmatrix}\right) = E\left(\begin{pmatrix} e_x^2 & e_x e_y \\ e_x e_y & e_y^2 \end{pmatrix}\right) = \\
&= E\left(\begin{pmatrix} e_x^2 & 2e_x^2 \\ 2e_x^2 & 4e_x^2 \end{pmatrix}\right) = \begin{pmatrix} \sigma_x^2 & 2\sigma_x^2 \\ 2\sigma_x^2 & 4\sigma_x^2 \end{pmatrix} = \sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\sigma_{x_1} &= \sigma_x \\
\sigma_{x_2} &= 2\,\sigma_x \\
\sigma_{x_1 x_2} &= 2\,\sigma_x^2 \\
\rho &= \frac{\sigma_{x_1 x_2}}{\sigma_{x_1}\,\sigma_{x_2}} = \frac{2\,\sigma_x^2}{\sigma_x\, 2\,\sigma_x} = 1
\end{aligned}
$$

Obviously, $\rho = 1$ because the relation between the random variables x and y is linear.
Calculate the eigenvalues and eigenvectors of the covariance matrix Σ.

$$
\begin{aligned}
\left|\Sigma - \lambda_j\, I\right| &= \left|\sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} \lambda_j & 0 \\ 0 & \lambda_j \end{pmatrix}\right| = \left|\begin{pmatrix} \sigma_x^2 - \lambda_j & 2\sigma_x^2 \\ 2\sigma_x^2 & 4\sigma_x^2 - \lambda_j \end{pmatrix}\right| = \\
&= \left(\sigma_x^2 - \lambda_j\right)\left(4\,\sigma_x^2 - \lambda_j\right) - 4\,\sigma_x^2 = -\lambda_j\,\sigma_x^2 - 4\,\sigma_x^2\,\lambda_j + \lambda_j^2 = \\
&= \lambda_j^2 - 5\,\sigma_x^2\lambda_j = \lambda_j\,\left(\lambda_j - 5\,\sigma_x^2\right) = 0
\end{aligned}
$$

$$
\begin{aligned}
\lambda_1 &= 0 \\
\lambda_2 &= 5\sigma_x^2
\end{aligned}
$$

$$
\left(\Sigma - \lambda_1\, I\right) \cdot \vec{u}_1 = \left(\sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}\right) \cdot \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} u_{11} \\ u_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

$$
\begin{aligned}
\sigma_x^2\, u_{11} + 2\,\sigma_x^2\, u_{12} &= 0 \\
2\,\sigma_x^2\, u_{11} + 4\,\sigma_x^2\, u_{12} &= 0 \longrightarrow \sigma_x^2\, u_{11} + 2\,\sigma_x^2\, u_{12} = 0
\end{aligned}
$$

$$
u_{11} = -\frac{2\,\sigma_x^2}{\sigma_x^2}\, u_{12} = -2\, u_{12}
$$

We choose:

$$
\begin{aligned}
u_{12} = 1 &\longrightarrow u_{11} = -2 \\
\vec{u}_1 &= \begin{pmatrix} -2 \\ +1 \end{pmatrix}
\end{aligned}
$$

Now, the vector $\vec{u}_1$ is normalized:

$$
\begin{aligned}
\|\vec{u}_1\| = \sqrt{\langle \vec{u}_1, \vec{u}_1 \rangle} = \sqrt{\vec{u}_1^T \cdot \vec{u}_1} = \sqrt{u_{11}^2 + u_{12}^2} = \sqrt{(-2)^2 + 1^2} = \sqrt{5} \\
\vec{u}_1 = \begin{pmatrix} \frac{u_{11}}{\|\vec{u}_1\|} \\ \frac{u_{12}}{\|\vec{u}_1\|} \end{pmatrix} = \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ +\frac{1}{\sqrt{5}} \end{pmatrix}
\end{aligned}
$$

$$\left(\Sigma - \lambda_2 I\right) \cdot \vec{u}_2 = \left(\sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} - \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2 \end{pmatrix}\right) \cdot \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \left(\sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} - 5\,\sigma_x^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \cdot \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} -4\,\sigma_x^2 & 2\,\sigma_x^2 \\ 2\,\sigma_x^2 & -\sigma_x^2 \end{pmatrix} \cdot \begin{pmatrix} u_{21} \\ u_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$-4\,\sigma_x^2\,u_{21} + 2\,\sigma_x^2\,u_{22} = 0 \longrightarrow 2\,\sigma_x^2\,u_{21} - \sigma_x^2\,u_{22} = 0$$
$$2\,\sigma_x^2\,u_{21} - \sigma_x^2\,u_{22} = 0$$

$$u_{21} = \frac{\sigma_x^2}{2\,\sigma_x^2}\,u_{22} = \frac{1}{2}\,u_{22}$$

We choose:

$$u_{22} = 2 \longrightarrow u_{21} = 1$$
$$\vec{u}_2 = \begin{pmatrix} +1 \\ +2 \end{pmatrix}$$

Now, the vector $\vec{u}_2$ is normalized:

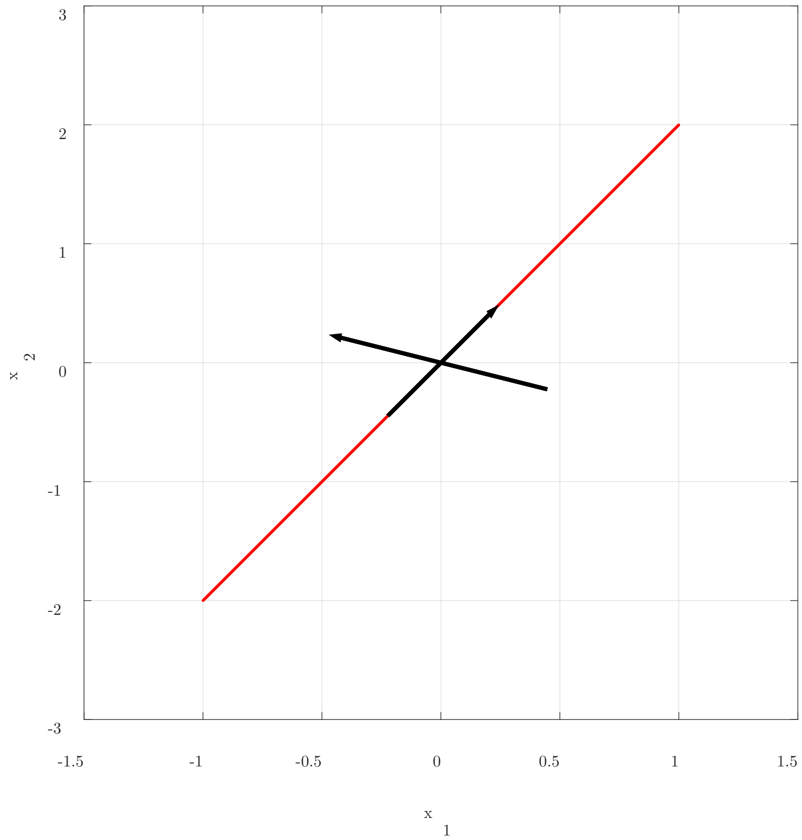
$$\|\vec{u}_2\| = \sqrt{\langle \vec{u}_2, \vec{u}_2 \rangle} = \sqrt{\vec{u}_2^T \cdot \vec{u}_2} = \sqrt{u_{21}^2 + u_{22}^2} = \sqrt{1^2 + 2^2} = \sqrt{5}$$
$$\vec{u}_2 = \begin{pmatrix} \frac{u_{21}}{\|\vec{u}_2\|} \\ \frac{u_{22}}{\|\vec{u}_2\|} \end{pmatrix} = \begin{pmatrix} +\frac{1}{\sqrt{5}} \\ +\frac{2}{\sqrt{5}} \end{pmatrix}$$

$$\vec{u} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Sigma = \sigma_x^2 \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} u_{11} & u_{21} \\ u_{12} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} =$$
$$= \begin{pmatrix} -\frac{2}{\sqrt{5}} & +\frac{1}{\sqrt{5}} \\ +\frac{1}{\sqrt{5}} & +\frac{2}{\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 5\,\sigma_x^2 \end{pmatrix} \cdot \begin{pmatrix} -\frac{2}{\sqrt{5}} & +\frac{1}{\sqrt{5}} \\ +\frac{1}{\sqrt{5}} & +\frac{2}{\sqrt{5}} \end{pmatrix}$$

Let's plot the ellipse associated to the covariance matrix Σ when $\sigma_x = 1$.

$$\Sigma = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} -\frac{2}{\sqrt{5}} & +\frac{1}{\sqrt{5}} \\ +\frac{1}{\sqrt{5}} & +\frac{2}{\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} \cdot \begin{pmatrix} -\frac{2}{\sqrt{5}} & +\frac{1}{\sqrt{5}} \\ +\frac{1}{\sqrt{5}} & +\frac{2}{\sqrt{5}} \end{pmatrix}$$



As we can see, because of the length of the ellipses' minor axis is 0, $\lambda_1 = 0$, the ellipse degenerates in a straight line. This result match perfectly with the fact I got earlier that the relation between the random variables x and y is linear, i.e, $y = 2\,x \longrightarrow \rho = 1$.