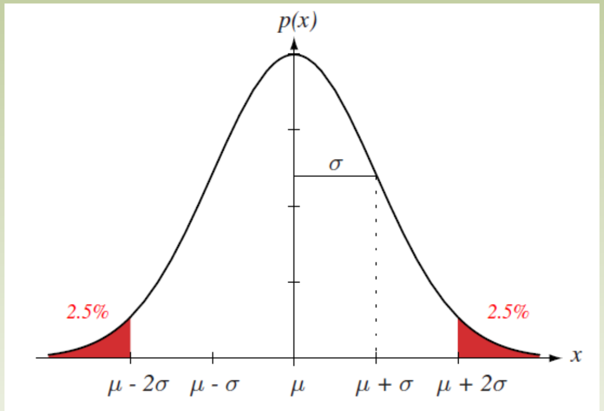
Multivariate Gaussian Distribution
(Multivariate Normal – MVN)

Univariate Gaussian

- Continuous univariate normal (Gaussian) probability density function:

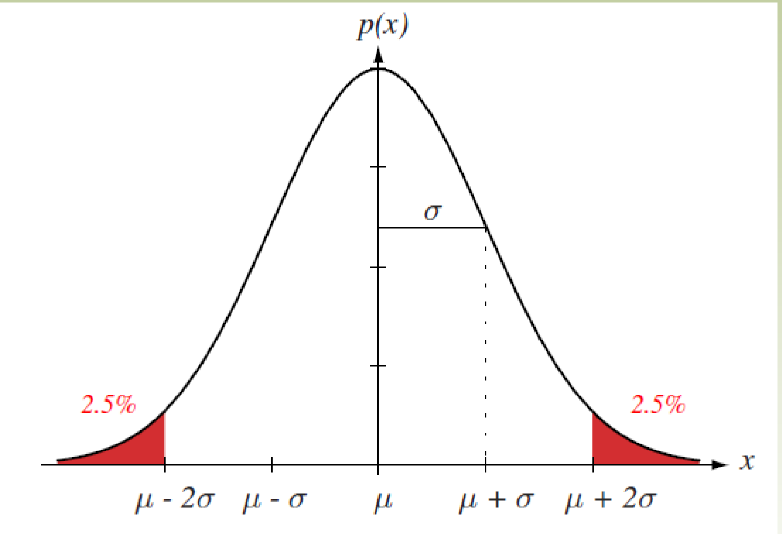
$$p(\mathbf{x}) \backsim \mathcal{N}(x|\mu, \sigma^2)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$



- The mean is the expected value of x is
    - $\mu = E[x] = \int_{-\infty}^{\infty} xp(x)dx$
- The variance is the squared of the standard deviation
    - $\sigma^2 = E[(x-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$

Sufficient Statistics

- Samples from the normal density tend to cluster around the mean and be spread-out based on the variance



- The normal density is completely specified by the mean and the variance. These are the <u>sufficient statistics</u>.

Multivariate Gaussian

$$p(\mathbf{x}) \backsim \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})]$$

- D : dimension
- $\mathbf{x}$: D-dimensional column vector
- $\boldsymbol{\mu}$: $\mathrm{E}[\mathbf{x}] \in \mathbb{R}^D$: D-dimensional mean vector
- Σ : covariance matrix (DxD)
- $|\Sigma|$: determinant of covariance matrix Σ
- Σ^{-1}: inverse of covariance matrix Σ

Univariate: $p(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$

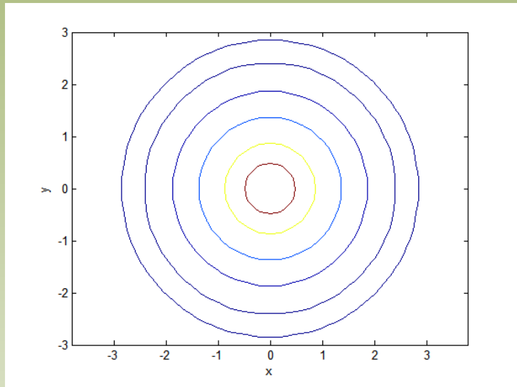
Covariance

- $\Sigma = \begin{bmatrix} var(x_1) & cov(x_1x_2) & cov(x_1x_3) \\ cov(x_2x_1) & var(x_2) & cov(x_2x_3) \\ cov(x_3x_1) & cov(x_3x_2) & var(x_3) \end{bmatrix}$

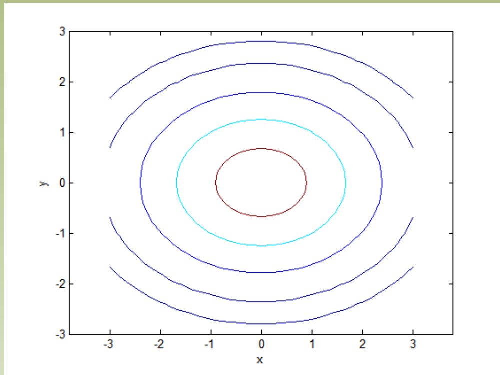
- Cov(X,Y) measures the degree in which X and Y are related
  - 0 → X,Y are statistically independent
  - > 0 → X,Y move in the same direction
  - < 0 → X,Y move in opposite direction

  - $cov(X,Y) = E[XY] - E[X]E[Y]$
  - $cov(X,X) = var(X)$
  - $cov(X,Y) = cov(Y,X)$

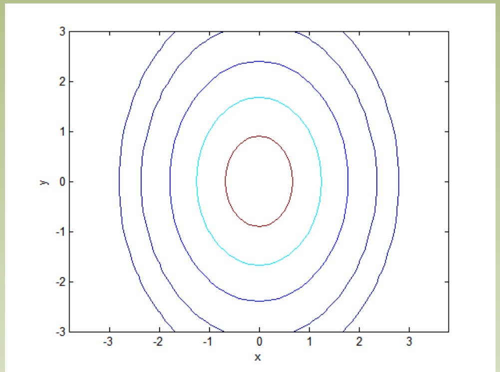
Level sets for 2D Gaussians – MVN_plot.m



$$\Sigma = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

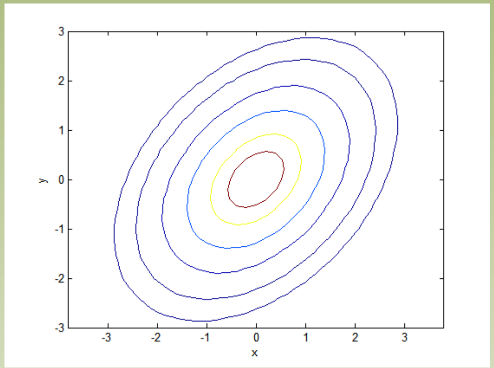


$$\Sigma = \begin{bmatrix} 0.9 & 0 \\ 0 & 0.5 \end{bmatrix}$$

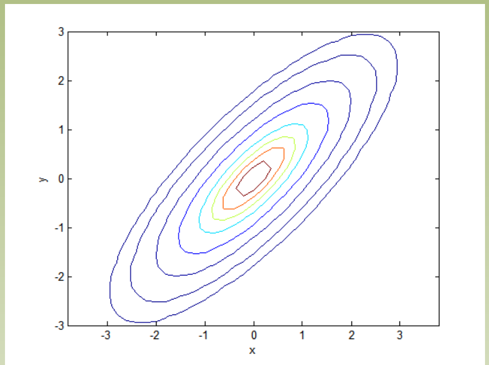


$$\Sigma = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.9 \end{bmatrix}$$

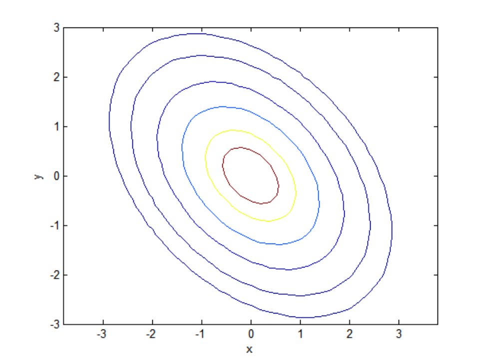
Level sets for 2D Gaussians – MVN_plot.m



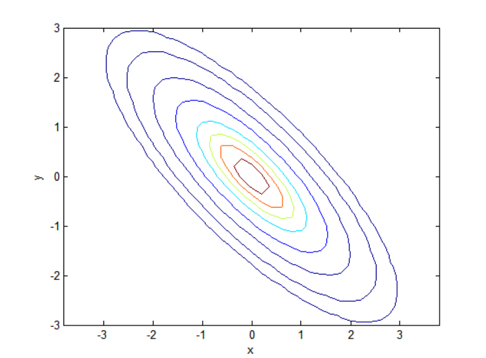
$$\Sigma = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 0.5 & 0.4 \\ 0.4 & 0.5 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 0.5 & -0.2 \\ -0.2 & 0.5 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 0.5 & -0.4 \\ -0.4 & 0.5 \end{bmatrix}$$

Mahalanobis Distance

- $r = \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}$

- Using eigen-decomposition, we can prove

  $r^2 = \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i}$ where $y_i = \mathbf{u}_i^T(\mathbf{x} - \boldsymbol{\mu})$

  λ_i : eigen values; $\mathbf{u}_i$: eigen vectors

- Principal axis
  - Eigenvectors determine the orientation
  - Eigenvalues determine the elongation