Further topics on multivariate continuous distributions

Dr. Yoann Altmann

B39AX – Fall 2023
Heriot-Watt University

Plan

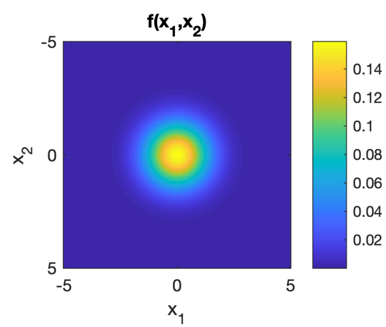
- Change of variable – 1D case
- Joint distributions
- Covariance and correlation
- Multivariate Gaussian distribution

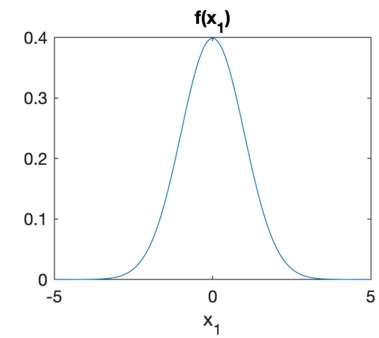
Introduction to the Multivariate Gaussian distribution

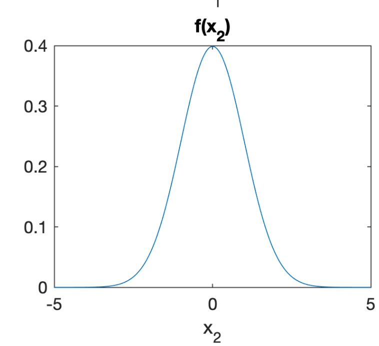
- Let X_1 and X_2 be centered i.i.d. Gaussian variables

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^2 e^{-\frac{x_1^2}{2\sigma^2}-\frac{x_2^2}{2\sigma^2}}$$

Example: $\sigma^2 = 1$







Introduction to the Multivariate Gaussian distribution

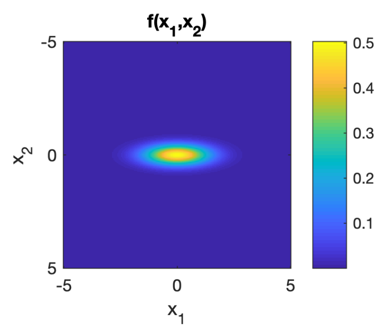
- If X_1 and X_2 are independent but non-identically distributed Gaussian variables

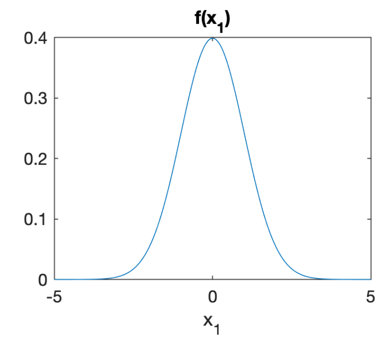
$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi\sigma_1^2}}e^{-\frac{x_1^2}{2\sigma_1^2}}\frac{1}{\sqrt{2\pi\sigma_2^2}}e^{-\frac{x_2^2}{2\sigma_2^2}}$$

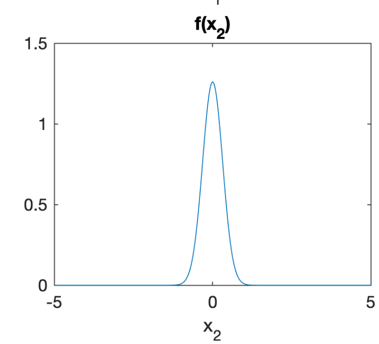
Example: $(\sigma_1^2, \sigma_2^2) = (1, 0.1)$

$$f_X(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^2 \det(D)^{1/2}\, e^{-\frac{x^T D^{-1} x}{2}}$$

With $D = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$ and $X = (X_1, X_2)$



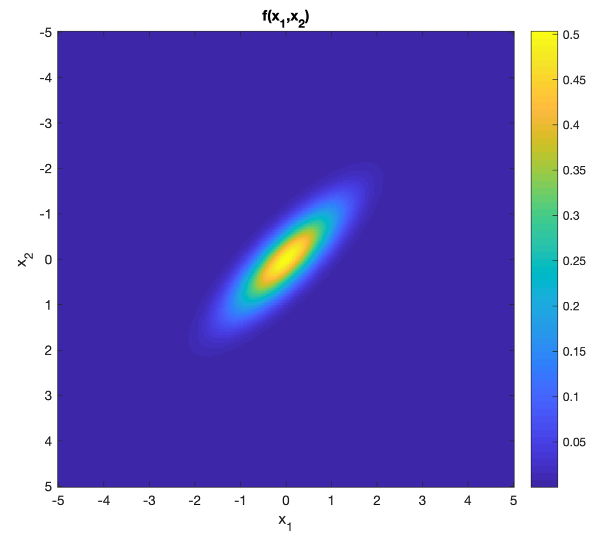




Introduction to the Multivariate Gaussian distribution

- How can we model correlation between Gaussian variables?
- Intuitively, the pdf seems

to be rotated …



- To investigate this, we need the notions of
  - Covariance/correlation
  - Change of variable (e.g., rotations)

Covariance and Correlation

- Covariance

Let X and Y be two random variables

$$cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$
$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

If X and Y are independent, $cov(X,Y) = 0$

⚠️ $cov(X,Y) = 0 \nrightarrow X$ and Y are independent

Example

- Let X and Y be two discrete random variables

(X,Y)	(1,0)	(0,1)	(-1,0)	(0,-1)
P(X,Y)	1/4	1/4	1/4	1/4

- $\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[XY] = 0$
- $cov(X, Y) = 0$ but X and Y are NOT independent ($\mathbb{p}(X|Y) \neq \mathbb{p}(X)$)
- X and Y are uncorrelated

Correlation coefficient

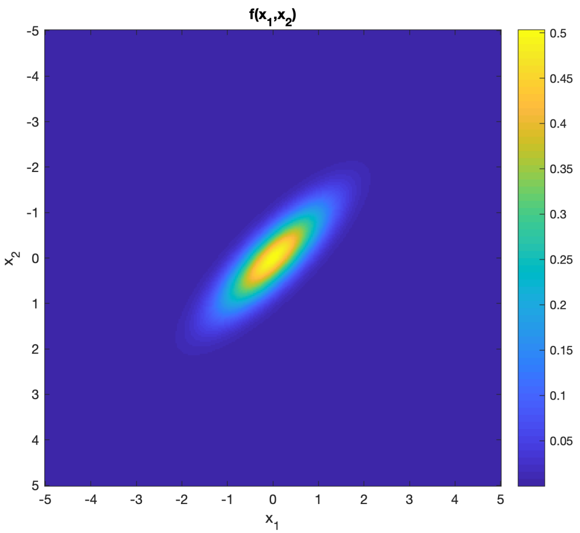
$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

- $\rho(X,Y) = 0$: X and Y are uncorrelated
- $\rho(X,Y) > 0$: X and Y are positively correlated
- $\rho(X,Y) < 0$: X and Y are negatively correlated

$$-1 \leq \rho(X,Y) \leq 1$$

Multivariate Gaussian distribution

- Here X_1 and X_2 are positively correlated
  - x_1 large (positive) implies x_2 likely to be large (using $f_{X_2|X_1}(x_2|x_1)$)



2D Gaussian distribution

Let $X = (X_1, X_2)$ be a 2D random vector following a multivariate Gaussian distribution with mean m and covariance matrix Σ

$$f_X(X) = \left(\frac{1}{\sqrt{2\pi}}\right)^2 det(\Sigma)^{-1/2}\, e^{-\frac{(x-m)^T\Sigma^{-1}(x-m)}{2}}$$

$$m = [m_1\ m_2]^{\mathrm{T}}, \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

$$\mathbb{E}[X] = m$$

2D Gaussian distribution

- Properties

$f_{X_1}(x_1) = \int f_X(x) dx_2$ is a Gaussian pdf

Marginal distributions:

$$X_1 \sim \mathcal{N}(x_1; m_1, \Sigma_{11}),\ X_2 \sim \mathcal{N}(x_2; m_2, \Sigma_{22}),$$

Conditional distributions:

$f_{X_1|X_2}(x_1|x_2) = \frac{f_X(x_1, x_2)}{\int f_X(x_1, x_2) dx_1}$ is a Gaussian pdf

2D Gaussian distribution

$f_{X_1|X_2}(x_1|x_2) = \frac{f_X(x_1,x_2)}{\int f_X(x_1,x_2)dx_1}$ is a Gaussian pdf

$$X_1|(X_2 = x_2) \sim \mathcal{N}(x_1; \widetilde{m}, \tilde{\Sigma})$$

$$\widetilde{m} = m_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - m_2)$$
$$\tilde{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Verification using Matlab

- See demo4.m

Multivariate Gaussian distribution

Let $X = (X_1, \ldots, X_N)$ be a N-D random vector following a multivariate Gaussian distribution with mean m and covariance matrix Σ

$$f_X(X) = \left(\frac{1}{\sqrt{2\pi}}\right)^N det(\Sigma)^{-1/2}\, e^{-\frac{(x-m)^T\Sigma^{-1}(x-m)}{2}}$$

- Σ is a symmetric positive-definite matrix.

$(x^T\Sigma x > 0, \forall x)$

- All the marginals and conditionals are (multivariate) Gaussian distributions

Covariance matrix

- Generalisation of variance for random vectors
  - 1-D case

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

  - N-D case (X is an $N \times 1$ vector)

$$COV(X) = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$$

⚠️ $COV(X)$ is an $N \times N$ matrix, not to be confused with the scalar

$$\mathbb{E}\left[(X - \mathbb{E}[X])^T (X - \mathbb{E}[X])\right]$$

Covariance matrix

- Multivariate Gaussian distribution

$$X \sim \mathcal{N}(x; m, \Sigma)$$

$$\mathbb{E}[X] = m$$
$$COV(X) = \mathbb{E}[(X - m)(X - m)^T] = \Sigma$$

Multivariate Gaussian distribution

Let X be an N-D random vector and Y be an M-D random vector. Let C be an $N \times M$ matrix.

If X and Y are independent and

$$X \sim \mathcal{N}(x; m_1, \Sigma_1)$$
$$Y \sim \mathcal{N}(y; m_2, \Sigma_2)$$

Then the random vector $Z = X + CY$ satisfies

$$Z \sim \mathcal{N}(z;\ \ m_1 + Cm_2, \Sigma_1 + C\,\Sigma_2 C^T)$$

Summary

- Change of variable – 1D case
- Joint distributions
- Covariance and correlation
- Multivariate Gaussian distribution

- Next chapter: Bayesian computation