Part 2b: Multivariate normal distributions

<u>Textbook</u>: pp. 41-45.

Introduction to the multivariate normal distribution

- The multinormal distribution is the "standard" statistical model for data.
- A data matrix contains variables as columns and observations as rows.
- Observations are independent, variables are not.
- According to the model, each observation has a multivariate normal distribution.
- Warning: The multivariate normal distribution is formulated as a column vector.

Definitions of multivariate normal distribution

Definition 2.1 Let $\mu \in \mathbb{R}^d$ be a vector and $\Sigma \in \mathbb{R}^{d \times d}$ be a symmetric, positive definite matrix. Then a random vector $X \in \mathbb{R}^d$ is normally distributed with mean μ and covariance matrix Σ, if the distribution of X has density $f : \mathbb{R}^d \to \mathbb{R}$ given by

$$f(x) = \frac{1}{(2\pi)^{d/2} \, |\det \Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1} (x-\mu) \right) \tag{2.1}$$

for all $x \in \mathbb{R}^d$.

An alternative definition

Lemma 2.2 Let $\mu \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times d}$ be ~~invertible~~. Define $\Sigma = AA^\top \in \mathbb{R}^d$. Furthermore, let $X = (X_1, X_2, \ldots, X_d) \in \mathbb{R}^d$ be a random vector such that $X_1, X_2, \ldots, X_d \sim \mathcal{N}(0, 1)$ are independent. Then

$$AX + \mu \sim \mathcal{N}(\mu, \Sigma)$$

on $\mathbb{R}^d$.

$$\vec{Y} = A\vec{X} + \vec{\mu} = g(\vec{X})$$

$$w(\vec{y}) = A^{-1}(\vec{y} - \vec{\mu})$$

$$J_w(\vec{y}) = A^{-1} \Rightarrow \det(J_w(y)) = \det(A^{-1}) = (\det(\Sigma))^{-\frac{1}{2}}$$

$$f_{\vec{X}}(\vec{x}) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2} = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}\vec{x}^T\vec{x}}$$

$$f_{\vec{Y}}(\vec{y}) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}[A^{-1}(\vec{y}-\mu)]^T[A^{-1}(\vec{y}-\vec{\mu})]} \cdot \frac{1}{\sqrt{\det(\Sigma)}}$$

$$= \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det(\Sigma)}} e^{-\frac{1}{2}(\vec{y}-\vec{\mu})^T\Sigma^{-1}(\vec{y}-\vec{\mu})}$$

Expectation and variance

Lemma 2.3 Let $X \sim \mathcal{N}(\mu, \Sigma)$ where $\mu \in \mathbb{R}^d$ and $\Sigma = (\sigma_{ij})_{i,j=1,\ldots,n} \in \mathbb{R}^{d \times d}$. Then

$$\mathbb{E}(X_i) = \mu_i$$

and

$$\operatorname{Cov}(X_i, X_j) = \sigma_{ij}$$

for all $i, j = 1, 2, \ldots, d$.

$$\mathbb{E}(\vec{X}) = \vec{0}, \quad \operatorname{Var}(\vec{X}) = I$$

$$\vec{Y} = A\vec{X} + \vec{\mu}$$

$$\mathbb{E}(\vec{Y}) = A\,\mathbb{E}(\vec{X}) + \vec{\mu} = \vec{\mu}$$

$$\operatorname{Var}(\vec{Y}) = A \operatorname{Var}(\vec{X}) A^T = AA^T = \Sigma$$

An example

Example 2.4 Let $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ be i.i.d. and define

$$X_i = \sum_{k=1}^{i} \varepsilon_k$$

$$\vec{X} = A\vec{\varepsilon}$$

$$A = \sigma^2 \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 1 & 1 & 1 & 1 & 1 \, 1 & 1 \end{pmatrix}$$

$$AA^T = \sigma^2 \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & 3 & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{pmatrix} = \Sigma$$

$$\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N(\vec{0}, \Sigma)$$