Lecture 6: Random Vectors

Professor Ilias Bilionis

The multivariate normal - diagonal covariance case

PREDICTIVE
SCIENCE LABORATORY

Multivariate normal - diagonal covariance case

- Take the special case of N independent random variables $X_1, \ldots, X_N$ each distributed according to a normal with known mean and variance:

$$X_i \sim N(\mu_i, \sigma_i^2)$$

- Seen as a random vector, the **joint pdf** of these variables is:

$$X = (X_1, \ldots, X_N)$$

because x_i's are independent

$$p(\mathbf{x}) = p(x_1, \ldots, x_N) = \prod_{i=1}^{N} p(x_i) = \prod_{i=1}^{N} N(x_i | \mu_i, \sigma_i^2)$$

$$= \prod_{i=1}^{N} (2\pi)^{-1/2} \sigma_i^{-1} \cdot \exp\left\{-\frac{1}{2} \frac{(x_i - \mu_i)^2}{\sigma_i^2}\right\}$$

PREDICTIVE SCIENCE LABORATORY

Multivariate normal - diagonal covariance case

- Take the special case of N independent random variables $X_1, \dots, X_N$ each distributed according to a normal with known mean and variance:

$$\boxed{X_i \sim N(\mu_i, \sigma_i^2)}$$

- Seen as a random vector, the **mean** of these variables is:

$$\underline{X} = (X_1, \dots, X_N)$$

$$\mathbb{E}[\mathbf{X}] = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_N] \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix} = \underline{\mu}$$

PREDICTIVE
SCIENCE LABORATORY

Multivariate normal - diagonal covariance case

- Take the special case of N independent random variables $X_1, \ldots, X_N$ each distributed according to a normal with known mean and variance:

$$X_i \perp X_j \quad , \text{if} \quad i \neq j$$

$$X_i \sim N(\mu_i, \sigma_i^2)$$

$$\mathbb{C}[X_i, X_j] = 0, \quad \mathbb{C}[X_i, X_i] = \mathbb{V}[X_i] = \sigma_i^2$$

- Seen as a random vector, the **covariance matrix** of these variables is:

$$\mathbb{C}[\mathbf{X}, \mathbf{X}] = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{bmatrix} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$$

PREDICTIVE SCIENCE LABORATORY

Multivariate normal - diagonal covariance case

- We say that the distribution of such a random vector is a **multivariate normal** with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)$.

- We write:

* bottom bar denotes vector

$$\mathbf{X} \sim N\left(\underline{\mu}, \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)\right)$$

$$p(\underline{x}) = (2\pi)^{-N/2} \underbrace{\left|\mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)\right|}_{\det(\mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2))}^{-1/2} \cdot \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)^{-1/2} \cdot (\underline{x}-\underline{\mu})\right\}$$

Isotropic covariance

$$p(x) = (2\pi)^{-N/2} \sigma^{-N} \cdot \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i - \mu_i)^2\right\}$$

- For the special case where all the variances are the same and equal to σ^2, we write:

$$\mathbf{X} \sim N(\underline{\mu}, \sigma^2 \underline{\underline{I}})$$

unit matrix in $\mathbb{R}^N$

$$\underline{\underline{I}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}$$

- In this special case the joint PDF is:

$$p(\mathbf{x}) = N(\underline{x} \mid \underline{\mu}, \sigma^2 \underline{\underline{I}})$$

$$= (2\pi)^{-N/2} \cdot \underbrace{\left|\sigma^2 \underline{\underline{I}}\right|^{-1/2}}_{\sigma^{-N}} \exp\left\{-\frac{1}{2}(\underline{x} - \underline{\mu})^T \underbrace{(\sigma^2 \underline{\underline{I}})^{-1}}_{\sigma^{-2}\underline{\underline{I}}} (\underline{x} - \underline{\mu})\right\}$$

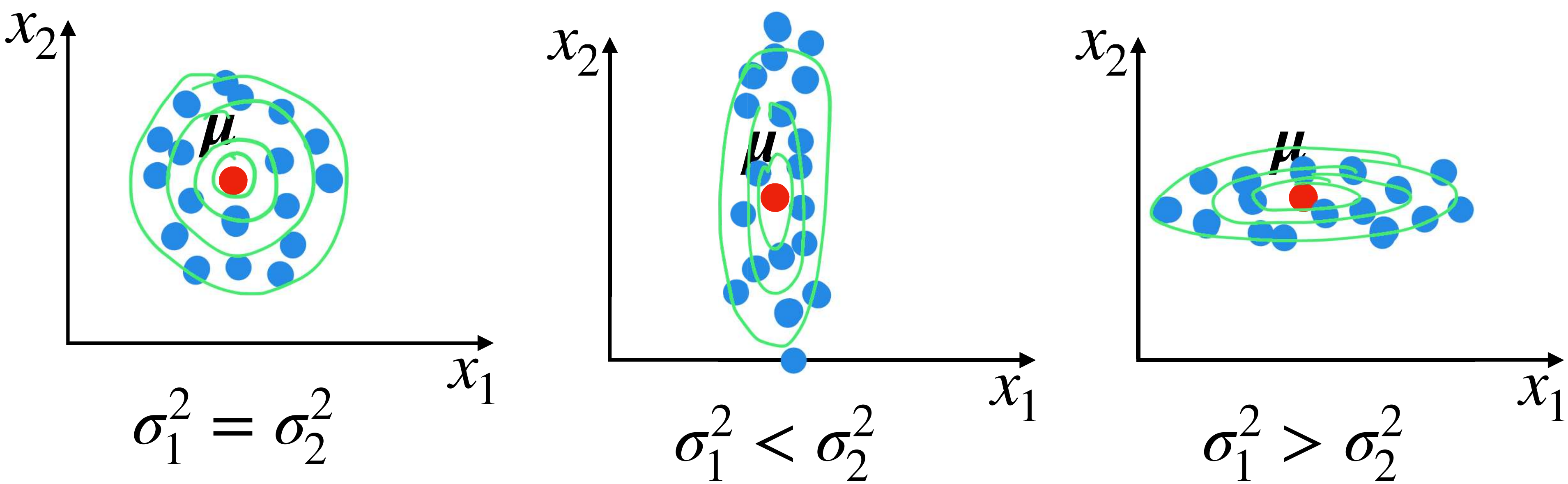
$$= \ldots$$

Visualizing the joint PDF of the multivariate normal with diagonal covariance

$X \sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma_1^2 & 0\\ 0 & \sigma_2^2\end{pmatrix}\right)$. $p(x) \propto \exp\left\{-\frac{1}{2\sigma_1^2}(x_1-\mu_1)^2 - \frac{1}{2\sigma_2^2}(x_2-\mu_2)^2\right\}$

proportional to

contour : a line $p(x) = \text{const}$



x_2 varies more

x_1 varies more

PREDICTIVE SCIENCE LABORATORY

Connection to the standard normal

- Let $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I})$ be a collection of independent standard normal random variables.

- Define the random vector:

$$\mathbf{X} = \boldsymbol{\mu} + \mathrm{diag}(\sigma_1, \ldots, \sigma_N)\mathbf{Z}$$

- Then:

$$\mathbf{X} \sim N\left(\underline{\mu}, \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2)\right)$$

PREDICTIVE SCIENCE LABORATORY