Gaussian random vectors

DS GA 1002 Probability and Statistics for Data Science

Carlos Fernandez-Granda

Prerequisites

Covariance matrix

Principal component analysis

Goals

Define Gaussian random vectors and explain connection to principal component analysis

Gaussian random variables

The pdf of a Gaussian or normal random variable $\tilde{a}$ with mean μ and standard deviation σ is given by

$$f_{\tilde{a}}(a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(a-\mu)^2}{2\sigma^2}}$$

Gaussian random vector

A Gaussian random vector $\tilde{x}$ is a random vector with joint pdf

$$f_{\tilde{x}}(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

where $\mu \in \mathbb{R}^d$ is the mean and $\Sigma \in \mathbb{R}^{d \times d}$ the covariance matrix

$\Sigma \in \mathbb{R}^{d \times d}$ is positive definite (positive eigenvalues)

Contour surfaces

Set of points at which pdf is constant when $\mu = 0$

$$
\begin{aligned}
c &= x^T \Sigma^{-1} x \\
&= x^T U \Lambda^{-1} U^T x \\
&= \sum_{i=1}^{d} \frac{(u_i^T x)^2}{\lambda_i}
\end{aligned}
$$

Ellipsoid with axes proportional to $\sqrt{\lambda_i}$

2D example

$$\mu = 0$$

$$\Sigma = \begin{bmatrix} 0.5 & -0.3 \\ -0.3 & 0.5 \end{bmatrix}$$

$$\lambda_1 = 0.8$$

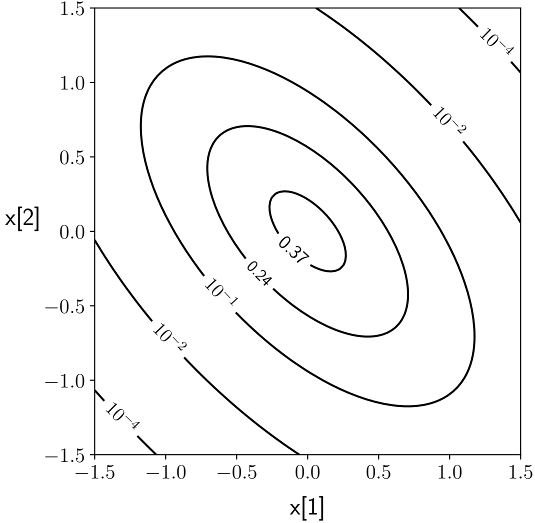
$$\lambda_2 = 0.2$$

$$u_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

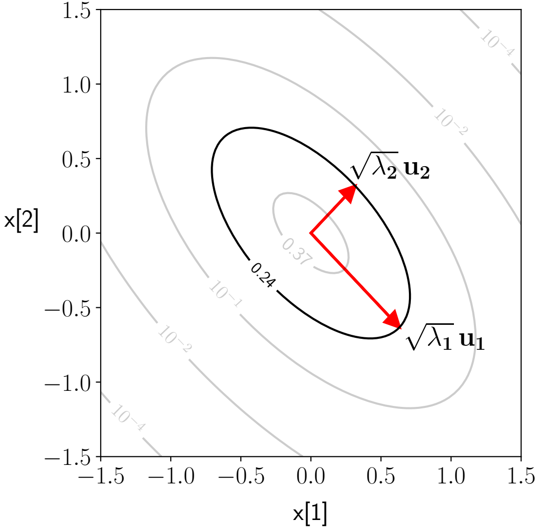
$$u_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

How does the ellipse look like?

Contour surfaces



Contour surfaces



Uncorrelation implies independence

If the covariance matrix is diagonal,

$$\Sigma_{\tilde{x}} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_d^2 \end{bmatrix}$$

the entries of a Gaussian random vector are independent

Proof

$$\Sigma_{\tilde{x}}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_d^2} \end{bmatrix}$$

$$|\Sigma| = \prod_{i=1}^{d} \sigma_i^2$$

Proof

$$
\begin{aligned}
f_{\tilde{x}}(x) &= \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) \\
&= \frac{1}{\prod_{i=1}^d \sqrt{(2\pi)}\sigma_i} \exp\left(\sum_{i=1}^d -\frac{(x[i]-\mu[i])^2}{2\sigma_i^2}\right) \\
&= \prod_{i=1}^d \frac{1}{\sqrt{(2\pi)}\sigma_i} \exp\left(-\frac{(x[i]-\mu[i])^2}{2\sigma_i^2}\right) \\
&= \prod_{i=1}^d f_{\tilde{x}[i]}(x[i])
\end{aligned}
$$

Linear transformations

Let $\tilde{x}$ be a Gaussian random vector of dimension d with mean μ and covariance matrix Σ

For any matrix $A \in \mathbb{R}^{m \times d}$ and $b \in \mathbb{R}^{m}$ $\tilde{y} = A\tilde{x} + b$ is Gaussian with mean $A\mu + b$ and covariance matrix $A\Sigma A^T$ (as long as it is full rank)

PCA on Gaussian random vectors

Let $\tilde{x}$ be a Gaussian random vector with covariance matrix $\Sigma := U \Lambda U^T$

The principal components

$$\widetilde{pc} := U^T \tilde{x}$$

are Gaussian and have covariance matrix

$$U^T \Sigma U^T = \Lambda$$

so they are independent

Only holds if data are Gaussian!

Maximum likelihood for Gaussian vectors

Log-likelihood of Gaussian parameters

$$(\mu_{\text{ML}}, \Sigma_{\text{ML}})$$

$$:= \arg \max_{\mu \in \mathbb{R}^d, \Sigma \in \mathbb{R}^{d \times d}} \log \prod_{i=1}^{n} \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2}(x_i - \mu)^T \Sigma^{-1} (x_i - \mu)\right)$$

$$= \arg \min_{\mu \in \mathbb{R}^d, \Sigma \in \mathbb{R}^{d \times d}} \sum_{i=1}^{n} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) + \frac{n}{2} \log |\Sigma| .$$

Solution is sample mean and variance

Additional justification, but PCA is useful without Gaussian assumption!

What have we learned

Definition of Gaussian random vectors and connection to principal component analysis