Machine Learning

Prof. Dr. Bernt Schiele
Dr. Paul Swoboda & Dr. Gerard Pons-Moll

Exercise Sheet 0 - 17.10.2018

This exercise does not have to be submitted. Please prepare the sheet for your next exercise group as it will be discussed.

Exercise 1 - Spectrum of Symmetric Matrices

Any real, symmetric matrix $A \in \mathbb{R}^{n\times n}$ has the decomposition

$$A = U\Sigma U^T,$$

where $\Sigma \in \mathbb{R}^{n\times n}$ is a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A on the diagonal and U is an orthogonal matrix in $\mathbb{R}^{n\times n}$, that is $UU^T = U^T U = \mathbb{1}$, which contains the corresponding eigenvectors (more precisely: an orthogonal basis of the eigenspace of the corresponding eigenvalue).

a. Derive the eigenvalues and eigenvectors of A^k (matrix product with itself) for $k \in \mathbb{N}$.

b. Prove that

$$\frac{\langle x, Ax\rangle}{\langle x, x\rangle} \leq \lambda_{\max}(A),$$

where $\langle x, y\rangle = x^T y$ is the inner product in $\mathbb{R}^n$, $\lambda_{\max}(A)$ is the largest eigenvalue of A.

Exercise 2 - Empirical Mean and Covariance

Given a set of n points $X = [x_1, \ldots, x_n]$, where $x_n \in \mathbb{R}^d, X \in \mathbb{R}^{d\times n}$.

a. Derive the minimizer c^* for function

$$f(c) = \sum_{i=1}^{n} \|x_i - c\|_2^2,$$

where $\|x\|_2$ denotes the Euclidean norm $\|c\|_2 = \sqrt{\sum_{j=1}^{d} c_j^2}$.

b. Show that the empirical covariance matrix for X

$$\Sigma_X = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)(x_i - \mu)^T$$

is positive semi-definite, that is $w^T \Sigma_X w \geq 0$, for all $w \in \mathbb{R}^d$, $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$.
Hint: consider using the Cauchy-Schwarz inequality, $\langle u, v\rangle^2 \leq \|u\|^2 \|v\|^2$.

Exercise 3 - Multivariate Gaussian

In the lecture we have seen the multivariate Gaussian $x \sim N(\mu, \Sigma)$ where the density function is defined as

$$f(x) = (2\pi)^{-\frac{d}{2}}|\Sigma|^{-\frac{1}{2}} exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu))$$

Now we have n multivariate Gaussian random vectors $\{x_i\}_{i=1}^n$, where $x_i \sim N(\mathbf{0}, \Sigma_i)$,$x_i \in \mathbb{R}^d$.

a. Consider the case where all the random vectors are mutually independent, derive the density function for $\sum_{i=1}^{n} x_i$.

b. Consider the case $n = 2$. Given the covariance matrix $cov(x_1, x_2) = C$, derive the density function for the joint vector $\binom{x_1}{x_2} \in \mathbb{R}^{2d}$.