Handed out: March 16 (Tue)

Hints for solution: Exercise class on March 19 (Fri)

Hand in: March 24 (Wed), by email to doris.entner@helsinki.fi. Please submit your report as a single PDF containing figures and discussion. Submit also the source code.

Solving the exercises below gives you points, which will, at the end, determine your grade for the computer project. Each exercise in all the assignments gives you an equal amount of points. In total, there will be approximately 15-20 exercises.

You can use the built-in commands of matlab or R to solve the exercises. In matlab, use `help` to get help on a command, e.g. `help eig`. In R, use `?` to get help on a command, e.g. `?eigen`.

Ex. 1 — Gaussian

1. Create $N = 5000$ realizations of a Gaussian random variable of mean $\mu = 5$ and variance $\sigma^2 = 2$.
2. Make a histogram of the counts in each bin of size $dx = 0.5$. Make also a histogram which shows an estimate of the gaussian *density*.
3. Calculate the sample mean $\bar{X}$ and sample variance S^2 (see math. exercise session 1).
4. Plug the estimates for the mean and variance into the formula for the Gaussian pdf and plot it in the same figure as the histogram from the previous question.

Ex. 2 — Eigenvalue Decomposition and Bivariate Gaussians

1. Create an unit vector $\mathbf{u}_1$ that forms angle α with the x-axis. Create vector $\mathbf{u}_2$ that is orthogonal to the first one. Put them into matrix U so that they are column vectors. Verify that $U^T U$ gives the identity matrix.
2. Create a diagonal matrix Λ with $\lambda_1 = 4$ and $\lambda_2 = 1$. Form $A = U\Lambda U^T$ and find then the eigenvectors and eigenvalues of A.
3. Create $N = 5000$ realizations of two independent Gaussian random variables $X = (X_1, X_2)$ with mean zero and variance 1. Create from X a random vector Y that has covariance matrix $\Sigma_y = A$. Check the result numerically. (*Friendly hint:* The covariance matrix Σ_y of $Y = MX$ is $M\Sigma_x M^T$ for some M.)
4. Do a scatter plot of Y, i.e. plot the data points of Y_1 versus the data points of Y_2.

Ex. 3 — Implementation of Gram-Schmidt

Implement the Gram-Schmidt method (treated in math exercise 1) to produce a set of orthonormal vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ from a set of (linearly independent) vectors $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. With normalization to unit norm, the Gram-Schmidt method was

$$\begin{aligned} \tilde{\mathbf{u}}_k &= \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{a}_\mathbf{k}^\mathbf{T} \mathbf{u}_\mathbf{i}) \mathbf{u}_i && (1) \\ \mathbf{u}_k &= \frac{\tilde{\mathbf{u}}_k}{||\tilde{\mathbf{u}}||_k} . && (2) \end{aligned}$$

Your program should take as input a matrix with the vectors $(\mathbf{a}_1, \ldots, \mathbf{a}_n)$, and output a matrix with the orthonormalized (orthogonal and unit norm) vectors $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$. Check your program with a simple example.