EECS 553 HW 1

Due Thursday, Sept. 5, by 11:59 PM Eastern Time

Gradescope: To submit this homework assignment, use Gradescope accessed through Canvas. To submit handwritten work, please first scan your assignment using a scanner or scanning app such as CamScanner. Please do not just submit a photo of your work as this will be much harder to read. Please allow some time in advance of the deadline to become familiar with Gradescope. When you upload your solutions to Gradescope, please indicate which pages of your solution are associated with each problem.

1. Honor Code (1 point each)

True or false: According to the Engineering Honor Code, as described in the Honor Code Pamphlet,

(a) It is the responsibility of faculty members to specify their policies in writing at the beginning of each semester. Students are responsible for understanding these policies and should consult the instructor if they are unclear.

(b) Students who are not members of the College of Engineering and who take a course offered by the College are bound by the policies of the Engineering Honor Code.

(c) If a student is accused of academic misconduct, they may simply withdraw from the class to avoid any blemish on their academic record.

2. PD/PSD matrices

(a) (3 points) Show that covariance matrices are positive semi-definite.

(b) (6 points) Let $\boldsymbol{A}$ be a symmetric matrix. Use the spectral theorem to show that $\boldsymbol{A}$ is invertible iff all of its eigenvalues are nonzero, and express the spectral decomposition of $\boldsymbol{A}^{-1}$ in terms of the spectral decomposition of $\boldsymbol{A}$. Conclude that if $\boldsymbol{A}$ is PD, then so is $\boldsymbol{A}^{-1}$.

3. Probability (3 points each)

Let X and Y be jointly distributed discrete random variables with joint probability mass function (pmf) $p(x, y)$. Prove the following using definitions and fundamental properties about random variables.

(a) Prove that for all x in the domain of X,

$$\Pr(X = x) = \sum_y p(x, y).$$

The quantity $p(x) := \Pr(X = x)$ is called the marginal pmf of X.
Note: Please do not use the law of total probability to prove this result. This avoids circular reasoning, because in an axiomatic development of probability, the result in 3 (a) is actually used to prove the law of total probability. Instead, just use the definition of joint pmf.

(b) Prove that $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$. Here, $\mathbb{E}[X] = \sum_x xp(x)$ denotes the expected value of X, and $\mathbb{E}[X|Y] = \sum_x xp(x|Y)$ denotes the conditional expectation of X given Y. Note that Y is random in this expression. Letting $p(y)$ denote the marginal distribution of Y, the entire right-hand side can be written $\sum_y (\sum_x xp(x|y))p(y)$.

4. Gaussian level sets (3 points each)

This problem merges probability and linear algebra. Let $\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, where $\boldsymbol{U}$ is the orthogonal matrix

$$\boldsymbol{U} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

$\theta = \frac{\pi}{6}$, and $\boldsymbol{\Lambda} = \mathrm{diag}(4, 3)$. Suppose that $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu} = [-1 \ \ -1]^T$.

(a) Sketch the boundary of

$$\mathcal{C} := \left\{ \boldsymbol{x} \,\middle|\, (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \leq r^2 \right\}$$

where $r = \sqrt{2}$. *Hint:* If we change $\leq r^2$ to $= r^2$, the set $\mathcal{C}$ is an ellipse. Your sketch should indicate the center, lengths of major and minor axes, and angle between the major axis and x-axis. Sketches may be hand drawn.

(b) Let $r = \sqrt{2}$. What is $\Pr(\boldsymbol{X} \in \mathcal{C})$? *Hint:* Use the chi-squared distribution and the following property of multivariate Gaussians: If $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and $\boldsymbol{A}$ is a matrix such that $\boldsymbol{A}\boldsymbol{X}$ is well-defined, then $\boldsymbol{A}\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{\mu}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T)$.

5. Unconstrained Optimization

(a) (3 points) Let $\boldsymbol{A}$ be an $m \times n$ matrix and $\boldsymbol{b} \in \mathbb{R}^m$. Consider a convex function $f : \mathbb{R}^m \to \mathbb{R}$. Using the definition of convexity, prove that $g(\boldsymbol{x}) = f(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$ is convex.

(b) (5 points) Consider the function $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$, where $\boldsymbol{A}$ is a symmetric $d \times d$ matrix. Derive the Hessian of f. Under what conditions on $\boldsymbol{A}$ is f convex? Strictly convex?