EECS 545 HW 1

Due Wednesday, Sept. 15, 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. Please allow some time in advance of the deadline to become familiar with Gradescope.

1. Honor Code

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. PSD matrices

Show that the following types of matrices are PSD:

(a) Any matrix of the form $\boldsymbol{A}^T\boldsymbol{A}$, where $\boldsymbol{A}$ is an arbitrary matrix.

(b) Covariance matrices, i.e., matrices of the form $\mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^T]$, where $\boldsymbol{X}$ is a random column vector.

(c) Gram matrices, i.e., any $d \times d$ matrix whose (i, j) entry is $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle$ for some vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$.

3. Probability

Let random variables X and Y be jointly distributed with distribution $p(x, y)$. You can assume that they are jointly discrete so that $p(x, y)$ is the probability mass function (pmf). Show the following results by using the fundamental properties of probability and random variables.

(a) $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$, where $\mathbb{E}[X] = \sum_x xp(x)$ denotes statistical expectation of X and $\mathbb{E}[X|Y] = \sum_x xp(x|Y)$ denotes conditional expectation of X given Y.

(b) $\mathbb{E}[I[X \in \mathcal{C}]] = P(X \in \mathcal{C})$, where $I[X \in \mathcal{C}]$ is the indicator function of an arbitrary set $\mathcal{C}$ (i.e. $I[X \in \mathcal{C}] = 1$ if $X \in \mathcal{C}$ and 0 otherwise.

(c) If X and Y are independent then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

(d) If X and Y take values in $\{0, 1\}$ and $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, then X and Y are independent.

4. Gaussian level sets

Let $\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$, where $\boldsymbol{U}$ is the orthogonal matrix

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

$\theta = -\pi/4$, and $\boldsymbol{\Lambda} = \mathrm{diag}(5, 1)$. Suppose that $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu} = [-2\ 3]^T$.

(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 $r = 1$. *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) 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) Use the definition of convexity to prove that the sum of two convex functions is convex.

(b) Use the Hessian to prove that the sum of two convex functions is convex. You may assume that the functions are twice continuously differentiable.

(c) 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?