EECS 127/227AT 2/15/2018
L. El Ghaoui

Quiz 1

NAME:

SID:

Instruction:

1. The quiz lasts 1h20.
2. The maximum score is 30.
3. Notes are *not* allowed, except for a one-page, two sided cheat sheet.
4. Do not open the exam until you are told to do so.

The breakdown of points is as follows.

Part	a	b	c	d	e	total
1	1	1	2	3		7
2	1	1	3	3		8
3	3	2	2	3	5	15

1. *Ellipse.* Consider the ellipse

$$\mathcal{E} = \left\{ x \;:\; (x - x_0)^\top P^{-1} (x - x_0) \le 1 \right\},$$

where

$$x_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^\top + \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}^\top .$$

Determine

(a) the center of the ellipse, $\hat{x}$;

(b) the semi-axes lengths, ρ_1, ρ_2;

(c) the corresponding principal directions u_1, u_2;

(d) an invertible matrix A and vector b such that in the coordinate system $\tilde{x} := Ax + b$, the ellipse looks like a sphere of radius 1 and center 0. (You need not provide the results in numerical form.)

Make sure to justify your answers carefully.

2. *Optimal weighting in a test.* A $n \times m$ matrix M contains the scores of n students on a test having m parts, so that M_{ij} is the score of student i in part j. We define a vector $w \in \mathbb{R}^m$ containing the weights w_j associated with each part $j = 1, \ldots, m$.

   (a) Express the vector $s \in \mathbb{R}^n$ containing the score of each student, in terms of M and w, in matrix notation.

   (b) Express the vector $\hat{m}$ containing the scores obtained in each part on average across students. You may denote by $m_i \in \mathbb{R}^m$ the i-th column of $M^\top$, $i = 1, \ldots, n$.

   (c) Someone suggests to the professor to use the maximum variance principle in order to compute a weight vector w. Explain how to do so, making sure to detail the covariance matrix involved.

   (d) What are possible shortcomings of the maximum-variance approach? *Hint:* comment on the sign of the entries of the maximum-variance weight vector w.

3. *PCA and optimal projection on a line.* In this exercise, we show the equivalence between PCA and a kind of least-squares problem involving a line.

   We consider a matrix of data points $X = [x_1, \ldots, x_m] \in \mathbb{R}^{n,m}$, and seek to find a line such that the sum of squared distances from the points to the line is minimized. In the sequel, we parametrize a generic line in $\mathbb{R}^n$ as

$$\mathcal{L}(x_0, u) = \{x_0 + tu \ : \ t \in \mathbb{R}^n\},$$

   where $x_0, u \in \mathbb{R}^n$ are given, with $\|u\|_2 = 1$. Geometrically, u provides the direction of the line, and x_0 its intercept.

   (a) Show that the distance from a given line $\mathcal{L}(x_0, u)$ to a given point $x \in \mathbb{R}^n$ is given by

$$D(x, \mathcal{L}(x_0, u))^2 = (x - x_0)^\top P(u)(x - x_0),$$

   where $P(u) := I_n - uu^\top$.

   (b) Is the symmetric matrix $P(u)$ positive semi-definite, definite? What are its eigenvalues?

   (c) What is the geometric interpretation of the linear map $x \to P(u)x$?

   (d) Now consider the minimization problem referred to above. Show that an optimal point x_0 is given by the center (average) of all data points. *Hint:* fix u and solve for x_0.

   (e) Show that an optimal direction u is given by the standard variance maximization problem at the heart of principal component analysis:

$$\max_u \ u^\top C u \ : \ \|u\|_2 = 1,$$

   where C is the covariance matrix of the data points.