EECS 127/227AT 9/25/2018
A. Bayen, L. El Ghaoui

Midterm

NAME:

SID:

Instruction:

1. The exam 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.
5. DO NOT WRITE ON THE BACK OF THE EXAM PAGES.

The breakdown of points is as follows.

Part	a	b	c	d	e	f	total
1	1	1	1	1			4
2	2	1	1	1	1	2	8
3	1	1	1	3	2	4	12
4	2	4	0	0	0	0	6

Name: ____________________

SID: ____________________

1.

Linear functions of images. In this problem we consider linear functions of an image with 2×2 pixels shown below.

3	7
8	5

This given image can be represented as the 4-vector $\begin{bmatrix} 3 \\ 7 \\ 8 \\ 5 \end{bmatrix}$.

Each of the operations described below defines a linear transformation $y = f(x)$, where the 4-vector x represents the original image, and the 4-vector y represents the resulting or transformed image. For each of these operations, provide the 4×4 matrix A for which $y = Ax$. **Also** in each case, determine the rank of the matrix A.

(a) Reflect the original image x across the vertical (i.e. bottom-to-top) axis.

(b) Rotate the original image x clockwise 90°.

Name: ______________________________

SID: ______________________________

(c) Rotate the original image x clockwise by $180°$.

(d) Set each pixel value y_i to be the average of the neighbors of pixel i in the original image. We define neighbors, to be the pixels immediately above and below and to the left and right. For the 2×2 matrix, every pixel has 2 neighbors.

Name: ______________________

SID: ______________________

2.

Fun with the SVD. Consider the 4×3 matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \tag{1}$$

where a_i for $i \in \{1, 2, 3\}$ form a set of *orthogonal* vectors satisfying $\|a_1\|_2 = 3$, $\|a_2\|_2 = 2$, $\|a_3\|_2 = 1$.

(a) What is the SVD of A? Express it as $A = USV^\top$, with S the diagonal matrix of singular values ordered in decreasing fashion, and explicitly describe U and V.

(b) Write A as a sum of 3 rank-one matrices.

Name: ______________________

SID: ______________________

(c) What is the dimension of the null space, $\dim(\text{null}(A))$?

(d) What is the rank of A, $\text{rank}(A)$? Provide an orthonormal basis for the range of A.

Name: ____________________

SID: ____________________

(e) Find the maximum "gain" of A (the amount that A can "expand" an input vectors ℓ_2 norm). More formally, what is the value of $\max_{x:\|x\|_2=1} \frac{\|Ax\|_2}{\|x\|_2}$?

(f) If I_3 denotes the 3×3 identity matrix, consider the matrix $\tilde{A} = \begin{bmatrix} A \\ I_3 \end{bmatrix} \in \mathbb{R}^{7\times 3}$? What are the singular values of $\tilde{A}$ (in terms of the singular values of A)?

Name: ________________________________

SID: ________________________________

3.

Regression and Applications. We first consider the regularized least-squares problem,

$$w_\lambda := \arg\min_w \|y - Xw\|_2^2 + \lambda\|w\|_2^2, \tag{2}$$

and subsequently an application to modeling time series. To begin, we investigate several fundamental properties of regression. Here, $X \in \mathbb{R}^{n,p}$ is the data matrix (with one data point per row), $y \in \mathbb{R}^n$ is the response vector, and $\lambda > 0$ is a "ridge" regularization parameter.

(a) Assume, only for this part, that $n < p$. Is $X^\top X$ invertible? Explain your reasoning.

(b) Now assume no relation between n and p. Is $X^\top X + \lambda I$ invertible? Explain your reasoning.

(c) Show that the solution to the full problem can be written as

$$w_\lambda = (X^\top X + \lambda I)^{-1} X^\top y \tag{3}$$

Name: ____________________

SID: ____________________

(d) Now suppose we would like find w that minimizes,

$$w = \arg\min_w \{\sum_{i=1}^{k} \lambda_i \|y_i - X_i w\|_2^2\} \tag{4}$$

(so it jointly fits k different linear regression objectives). Explain how to reformulate this problem as a *single* least-squares problem with augmented $\tilde{X}$ and $\tilde{y}$ in an objective $\|\tilde{y} - \tilde{X}w\|_2^2$, *and* find the solution w to the aforementioned problem[1].

(e) What is the computational complexity (in big-O notation) of computing the solution the previous question in terms of n, p, k? Assume each $X_i \in \mathbb{R}^{n \times p}$ and $y_i \in \mathbb{R}^n$ and once again that the relevant square matrices are invertible.

[1]you may assume the relevant square matrices are invertible.

Name: ____________________________

SID: ____________________________

Periodic Time Series. We now consider an application to the problem of modeling a periodic time series z_t which we approximate by a sum of K sinusoids:

$$z_t \approx \hat{z}_t = \sum_{k=1}^{K} a_k \cos(\omega_k t - \phi_k) \quad t = 1, 2, \ldots, T \tag{5}$$

The coefficient $a_k \geq 0$ are the amplitudes, ω_k the frequencies, and ϕ_k the phases. In many applications (and the one we consider) the frequencies ω_k are apriori known and **fixed**. We wish to find $a_1, \ldots, a_K$ and $\phi_1, \ldots, \phi_K$ to ensure the means-squared value of the approximation error $(\hat{z}_1 - z_1, \ldots, \hat{z}_T - z_t)$ is small.

(f) Explain how to solve the aforementioned problem using (regularized) least squares to estimate $a_1, \ldots, a_K$ and $\phi_1, \ldots, \phi_K$. Be explicit in the mappings between the values $z_t, a_k, \omega_k, \phi_k$ in the original formulation and the standard regression parametrization y, X, w_λ (detailed in the beginning of the question), and the dimensions of the relevant vectors/matrices. *Hint: Recall the identity* $a\cos(\omega t - \phi) = \alpha \cos(\omega t) + \beta \sin(\omega t)$ *for* $\alpha = a\cos\phi$ *and* $\beta = a \sin\phi$, *with* $a = \sqrt{\alpha^2 + \beta^2}$ *and* $\phi = \arctan(\beta/\alpha)$.

Name: ________________________

SID: ________________________

4.

Positive-Definite Matrices and Hessians.

Let $C \in \mathbb{R}^{n \times n}$ by a real, symmetric positive-definite matrix. Consider the function

$$f_\lambda(x) = \|C - xx^\top\|_F^2 + 2\lambda\|x\|_2^2$$

where $x \in \mathbb{R}^n$.

(a) Compute the Hessian matrix of the function $f_\lambda(x)$ with respect to x, $\nabla^2 f_\lambda(x)$.
Hint: Note that $\|C - xx^\top\|_F^2 = \|C\|_F^2 + \|x\|_2^4 - 2x^\top Cx$.

Name: ____________________________

SID: ____________________________

(b) When is the Hessian matrix (which depends on x), positive semi-definite at *all* points x? You should derive an "if and only if" condition expressed in terms of λ and a function of the matrix C.