Midterm Examination

CS 205A: Mathematical Methods for Robotics, Vision, and Graphics (Spring 2015),
Stanford University

- The exam runs for 75 minutes.
- The exam contains six problems. You must complete the first problem and four of problems 2-6. **CIRCLE THE PROBLEMS YOU WANT GRADED ON THE CHART BELOW; OTHERWISE WE WILL GRADE THE FIRST FOUR QUESTIONS ON WHICH YOU HAVE PROVIDED ANY WRITTEN ANSWER.**
- The exam is closed book/notes. You may use one double-sided $8^1/_2'' \times 11''$ sheet of notes.
- Write your solutions in the space provided. If you need more space, write on the back of the sheet containing the problem and indicate that you have done so.
- Do not spend too much time on any problem. Read them all before beginning.
- Show your work, as partial credit will be awarded.

Circle the four additional problems you want graded.

Problem	(1)	2	3	4	5	6	Total
Score (out of 10)							

The Stanford Honor Code

1. The Honor Code is an undertaking of the students, individually and collectively:

   (a) that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;

   (b) that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

2. The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.

3. While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

Signature

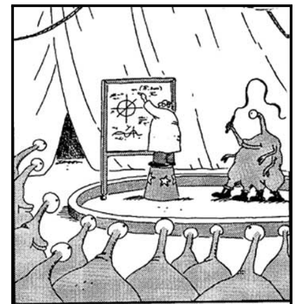
Name

SUID

YOU MUST COMPLETE THIS PROBLEM.

Problem 1 (Short answer).

(a) Derive a nontrivial linear system of equations satisfied by minima of $f(\vec{x}) \equiv \|A\vec{x} - \vec{b}\|_2$ with respect to $\vec{x}$, given $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^m$. [2 points]



Abducted by an alien circus company, Professor Winters is forced to write Linear Algebra equations in center ring.

(b) Consider the design of a piece of numerical software for simulating weather patterns. Provide an example of each of the following sources of potential error:

(i) Rounding error [1 point]:

(ii) Discretization error [1 point]:

(iii) Modeling error [1 point]:

(iv) Input error [1 point]:

(c) Suppose $A \in \mathbb{R}^{n\times n}$ satisfies $A^\top = A$. Explain what it means to *diagonalize* A; your formula should not involve a matrix inverse. [2 points]

CIRCLE WHICH FOUR OF PROBLEMS 2-6 YOU WANT GRADED ON PAGE 1.

Problem 2 (Least-squares).

(a) Suppose $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times k}$ are given, where $m \geq n$ and A is full-rank. Propose an algorithm for finding a matrix $X \in \mathbb{R}^{n\times k}$ solving the following variational problem:

$$\min_{X\in\mathbb{R}^{n\times k}} \|AX - B\|_{\text{Fro}}.$$

For full credit, the runtime of your algorithm should scale like $O(n^2k)$ when $k \gg m$. [5 points]

(b) The *variance* of a set of values $y_1, \ldots, y_k$ stored in a vector $\vec{y} \in \mathbb{R}^k$ is

$$\text{Var}[\vec{y}] \equiv \frac{1}{k}\sum_{i=1}^{k}(y_i - E[\vec{y}])^2,$$

where $E[\vec{y}]$ is the *mean* of $\vec{y}$: $E[\vec{y}] \equiv \frac{1}{k}\sum_{i=1}^{k} y_i$. Show how to solve the following optimization problem using linear techniques:

$$\min_{\vec{y}\in\mathbb{R}^k} \left(\|A\vec{y} - \vec{b}\|_2^2 + \gamma\text{Var}[\vec{y}]\right).$$

You can assume $A \in \mathbb{R}^{n\times k}$ is tall and full-rank and that $\gamma > 0$. [5 points]

CIRCLE WHICH FOUR OF PROBLEMS 2-6 YOU WANT GRADED ON PAGE 1.

Problem 3 (SVD).

(a) Find the singular value decomposition of the *inverse* of the matrix [3 points]

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 6 \\ 1 & 0 & 0 \end{pmatrix}.$$

(b) The remainder of this problem deals with computing the SVD of a matrix $A \in \mathbb{R}^{m\times n}$.

(i) Suppose for any matrix $B \in \mathbb{R}^{m\times n}$ we have a "black box" function that can compute

$$\vec{v}_{\max}(B) = \arg \max_{\|\vec{v}\|_2=1} \|B\vec{v}\|_2.$$

Using this function, describe an algorithm to compute the right singular vectors and singular values of A. [4 points]

(ii) Suppose instead we have a "black box" for computing

$$\sigma_{\max}(B) = \max_{\|\vec{v}\|_2=1} \|B\vec{v}\|_2.$$

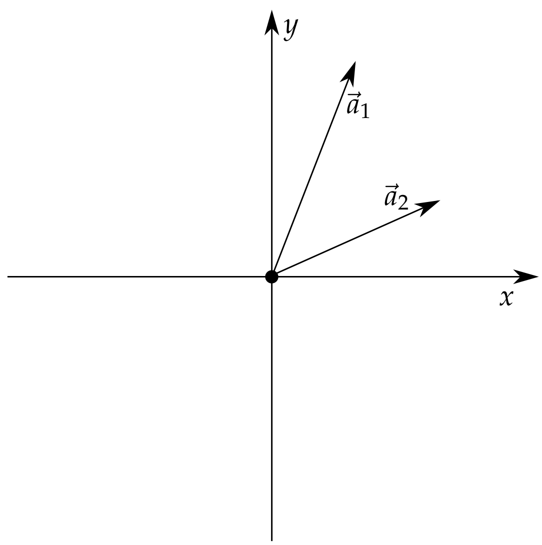
Can you still compute the right singular vectors of A? How? [3 points]

Note: You can propose using Gaussian elimination, but not eigenvalue computation.

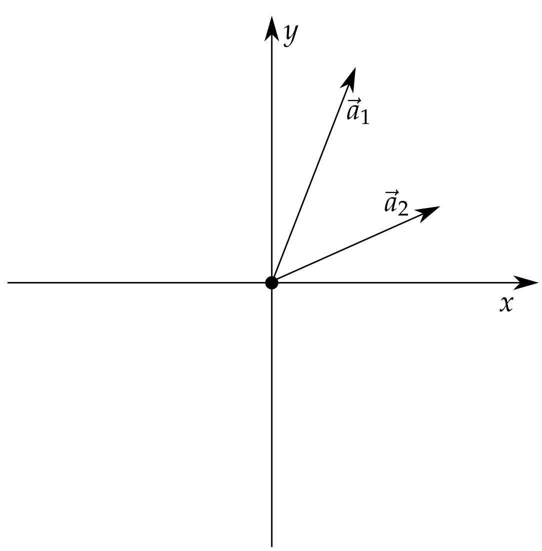
CIRCLE WHICH FOUR OF PROBLEMS 2-6 YOU WANT GRADED ON PAGE 1.

Problem 4 (QR and orthogonalization).

(a) In the following diagram, draw the set of orthonormal vectors that would result after applying Gram-Schmidt orthogonalization to the vectors $\vec{a}_1$ and $\vec{a}_2$, assuming that $\|\vec{a}_1\|_2 \approx 2$; label them as $\vec{b}_1$ and $\vec{b}_2$. [3 points]



(b) In the following diagram, draw the line across which the first step of Householder orthogonalization would reflect vector $\vec{a}_1$. Also, draw the images of $\vec{a}_1$ and $\vec{a}_2$ under this reflection; label them as $\vec{b}_1$ and $\vec{b}_2$. [3 points]



PROBLEM 4 IS CONTINUED ON NEXT PAGE.

(c) Consider the Vandermonde matrix

$$A = \begin{pmatrix} 1 & -0.99 & (-0.99)^2 \\ 1 & -0.98 & (-0.98)^2 \\ \vdots & \vdots & \vdots \\ 1 & 0.98 & (0.98)^2 \\ 1 & 0.99 & (0.99)^2 \end{pmatrix} \in \mathbb{R}^{199\times 3}.$$

Let $p(x) = ax^2 + bx + c$ for some arbitrary coefficients $a, b, c \in \mathbb{R}$, and let $\vec{p}$ represent the vector of $p(x)$ sampled uniformly on the interval $(-1, 1)$:

$$\vec{p} \equiv \begin{pmatrix} p(-0.99) \\ p(-0.98) \\ \vdots \\ p(0.98) \\ p(0.99) \end{pmatrix}.$$

Suppose we factor $A = QR$, with $Q \in \mathbb{R}^{199\times 3}$ and $R \in \mathbb{R}^{3\times 3}$, where the columns of Q are $\vec{q}_1$, $\vec{q}_2$, and $\vec{q}_3$. Given the relationships $\vec{p}^T\vec{q}_1 = 3$, $\vec{p}^T\vec{q}_2 = 2$, and $\vec{p}^T\vec{q}_3 = 6$, provide a numerical estimate of the quantity [4 points]

$$\sqrt{\int_{-1}^{1} p^2(x)dx}.$$

Note: You need not simplify your answer.

CIRCLE WHICH FOUR OF PROBLEMS 2-6 YOU WANT GRADED ON PAGE 1.

Problem 5 (Eigenproblems). For a square matrix $B \in \mathbb{R}^{n\times n}$, define the *off-diagonality criterion* as

$$\mathrm{off}(B) \equiv \sum_{i=1}^{n} \sum_{j\neq i} B_{ij}^2.$$

This objective function appears in methods for transfer learning and surface mapping.

(a) Suppose $A^\top = A$. Then, provide a solution to the following optimization problem: [4 points]

$$\min_{\substack{U\in\mathbb{R}^{n\times n}\\ U^\top U = I_{n\times n}}} \mathrm{off}(U^\top A U).$$

(b) In homework 3, you showed that QR factorization can be computed by expressing Q as a product of *Givens rotations* $G(i,j,\theta)$. Given an orthogonal matrix $U \in \mathbb{R}^{n\times n}$, argue there exists a diagonal matrix $D \in \{-1,0,1\}^{n\times n}$ so that UD is a product of Givens rotations. [3 points]

(c) Given symmetric matrices $A_1, \ldots, A_k \in \mathbb{R}^{n\times n}$, *simultaneous diagonalization* algorithms attempt to solve the following optimization problem (not equivalent to computing eigenvalues!):

$$\min_{\substack{U\in\mathbb{R}^{n\times n}\\ U^\top U = I_{n\times n}}} \sum_{\ell=1}^{k} \mathrm{off}(U^\top A_\ell U).$$

For fixed $i, j \in \{1, \ldots, n\}$, provide a one-variable nonlinear optimization problem for updating an estimate U to an improved estimate $\overline{U} = G(i,j,\theta)U$. [3 points]

Note: This is the building block of the "JADE" algorithm for simultaneous diagonalization, built on the observation from part (b).

CIRCLE WHICH FOUR OF PROBLEMS 2-6 YOU WANT GRADED ON PAGE 1.

Problem 6 (Matrix factorization). We can use Cholesky factorization to calibrate a camera by computing a matrix $W \in \mathbb{R}^{3\times 3}$ constructed using data collected from a single image. Assume W is symmetric positive definite with Cholesky factorization $W = LL^\top$.

(a) Is W necessarily invertible? Explain your answer. [2 points]

(b) We can construct a linear system to solve for the elements of W. To account for scaling, we fix $w_{33} = 1$. If the camera has no skew, we fix $w_{12} = 0$, and if pixels are square we take $w_{11} = w_{22}$. How many additional linear relationships are needed to find W? [2 points]

Note: Do not worry about constraining W to be positive definite.

(c) Prove that the factorization $W = LL^\top$ is unique. That is, given factorizations $W = LL^\top$ and $W = \tilde{L}\tilde{L}^\top$ where L and $\tilde{L}$ are lower triangular, prove that $L = \tilde{L}$. Assume L and $\tilde{L}$ have positive entries along their diagonals. [4 points]

Hint: Manipulate the relationship $LL^\top = \tilde{L}\tilde{L}^\top$ so that one side is lower triangular and the other is upper triangular.

(d) Now, consider a new symmetric matrix $M \in \mathbb{R}^{3\times 3}$, but assume there exists $\vec{x} \in \mathbb{R}^3$ with $\vec{x}^\top M\vec{x} < 0$. Can we apply Cholesky factorization to M? Why? [2 points]