Final Examination

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

- The exam runs for 3 hours.
- The exam contains seven problems. You must complete the first problem and five of problems 2-7. **CIRCLE THE PROBLEMS YOU WANT GRADED ON THE CHART BELOW; OTHERWISE WE WILL GRADE THE FIRST FIVE QUESTIONS ON WHICH YOU HAVE PROVIDED ANY WRITTEN ANSWER.**
- The exam is closed-book. You may use two double-sided $8^1/_2'' \times 11''$ sheets of notes.
- Write your solutions in the space provided. If you need more space, write on the back of the sheet and indicate that you have done so. Additional pages are provided at the back; please indicate on the original problem if you continue your answer there.
- 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 five additional problems you want graded.

Problem	1	2	3	4	5	6	7	EC	Total
Score									

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) Suppose we approximate the derivative of $f : \mathbb{R} \to \mathbb{R}$ using divided differences:

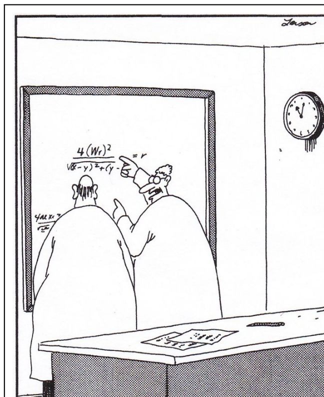
$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

(i) What can go wrong if h is too large? [1 point]

(ii) What can go wrong if h is too small? [1 point]

(b) Given full-rank $A, B \in \mathbb{R}^{n \times n}$ and $\vec{c} \in \mathbb{R}^n$, provide an algorithm for minimizing $f(\vec{x}, \vec{y}) = \|A\vec{x} + B\vec{y} - \vec{c}\|_2$ with respect to both $\vec{x}$ and $\vec{y}$. [2 points]

(c) Suppose we are given five data points as (x, y) pairs: $(-2, -4), (-1, 7), (0, 2), (2, 1/2), (5, 9)$. Compute $f(3) - g(3)$, where $f(\cdot)$ interpolates the data using the degree-four Lagrange basis and $g(\cdot)$ interpolates the data in the degree-four Newton basis. [2 points]



"Yes, yes, I *know* that, Sidney ... *everybody* knows *that!* ... But look: Four wrongs *squared*, minus two wrongs to the fourth power, divided by this formula, *do* make a right."

YOU MUST COMPLETE THIS PROBLEM.

(d) For a given differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, give the forward Euler iteration for solving the ODE $\vec{x}' = -\nabla f(\vec{x})$ with time step $h > 0$. Provide an alternative name for the resulting iterative scheme. [2 points]

(e) Can you read off the eigenvalues of any matrix $A \in \mathbb{R}^{n \times n}$ from its SVD? Explain how or provide a counterexample. [2 points]

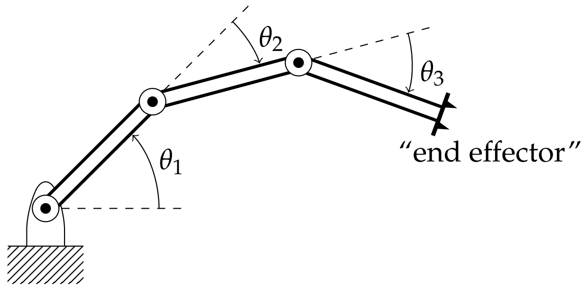


Figure 1: Planar robot arm for $n = 3$

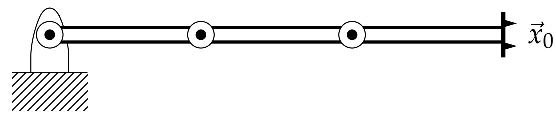


Figure 2: Resting configuration

Problem 2 (Robotics). A planar robot arm consists of n links connected by n joints, with an angle θ_i specified at each joint (Figure 1). Let $\vec{\theta} \in \mathbb{R}^n$ denote a vector of all the joint angles.

The tip of the robot arm is called its *end effector*. The position $\vec{x}_1 \in \mathbb{R}^2$ of the end effector is related to $\vec{\theta}$ by $\vec{x}_1 = K(\vec{\theta})\vec{x}_0$, where $\vec{x}_0$ is the position of the end effector in resting configuration (Figure 2) and $K(\vec{\theta})$ is a 2×2 "kinematic matrix" whose components are functions of $\vec{\theta}$.

(a) Suppose the position of the end effector is fixed at $\vec{x}_1 \in \mathbb{R}^2$. The potential energy of link ℓ is proportional to $h_\ell(\vec{\theta})$, where $h_\ell : \mathbb{R}^n \to \mathbb{R}$ is the height of the center of mass for link ℓ measured from the flat surface on which the arm is mounted. The arm should *never* go through the surface. To avoid singular configurations, we also constrain each θ_i to be in an interval $[a_i, b_i]$.

Write an optimization problem to find the minimum potential energy configuration of the manipulator, subject to the given constraints. [5 points]
NOTE: You may leave your objective in terms of the functions $\{h_\ell(\vec{\theta})\}_{\ell=1}^n$.

PROBLEM 2 IS CONTINUED ON THE NEXT PAGE.

(b) Write the KKT conditions for your optimization problem. [5 points]

CIRCLE WHICH SIX OF PROBLEMS 2-7 YOU WANT GRADED ON PAGE 1.

Problem 3 (Inner product matrix). For $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in \mathbb{R}^n$, define $M \in \mathbb{R}^{k\times k}$ to be the pairwise dot product matrix, with elements $M_{ij} \equiv \vec{x}_i \cdot \vec{x}_j$. This matrix is the focus of algorithms in metric embedding and learning that deduce geometric structures from weak observations about shape.

(a) Show that M is symmetric and positive semidefinite. [2 points]

(b) For $i, j \in \{1, \ldots, k\}$, show how to compute $\|\vec{x}_i - \vec{x}_j\|_2$ from elements of M. [1 point]

(c) Using your answer to the previous part, suggest how to use a linear solve to compute M given a matrix $D \in \mathbb{R}^{k\times k}$ with elements $D_{ij} \equiv \|\vec{x}_i - \vec{x}_j\|_2$ and a vector $\vec{v}$ with elements $v_i \equiv \|\vec{x}_i\|_2$. [3 points]
NOTE: You do not need to show that your matrix is invertible.

(d) Suppose $k \geq n$. Given such a matrix M, propose a method for recovering vectors $\vec{x}_1, \ldots, \vec{x}_k$ whose dot products are the elements of M. [4 points]

Problem 4 (ODE). Suppose a particle's velocity is given by the dynamical system

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

where $a(t)$, $b(t)$, $c(t)$, and $d(t)$ are functions of time. We wish to approximate $(x(t), y(t))$ numerically, yielding a sequence of values $(x_k, y_k) \approx (x(kh), y(kh))$ for some time step $h > 0$.

(a) Suppose a, b, c, and d are constant in time. Give an expression for computing (x_{k+1}, y_{k+1}) from (x_k, y_k) using trapezoidal integration. [2 points]

(b) Is the ODE stable if $a(t) = 2.5$, $b(t) = -6.0$, $c(t) = -6.0$, and $d(t) = -2.5$? Why? [3 points] NOTE: The printed text made an erroneous remark on stability in this case that was repaired in lecture and in the online version.

(c) Now, consider taking $a(t) = -4 - t$, $b(t) = -3t$, $c(t) = 0$, and $d(t) = t - 5$. Provide an expression for integrating this system in time using backward Euler. [3 points]

(d) Conjecture the range of values of t for which the ODE and/or its integrator from part (c) is stable. [2 points]

Problem 5 (Orthogonal matching pursuit). "Sparse signal recovery" algorithms approximate $\vec{v} \in \mathbb{R}^n$ as a linear combination of m columns of a matrix $\Phi \in \mathbb{R}^{n\times d}$, where $m \ll d$. A greedy algorithm for this task called *orthogonal matching pursuit* (Tropp & Gilbert 2007) is documented below:

1. Initialize residual $\vec{r} \leftarrow \vec{v}$ and iteration counter $t \leftarrow 1$. Initialize Φ_0 as an empty matrix.
2. Compute $\ell \leftarrow \arg\max_{j\in\{1,\ldots,d\}} |\vec{r}\cdot\vec{\varphi}_j|$, where $\vec{\varphi}_j$ is the j-th column of Φ.
3. Add $\vec{\varphi}_\ell$ to Φ_{t-1} as its rightmost column to compute Φ_t, that is, $\Phi_t \leftarrow (\Phi_{t-1}\ \ \vec{\varphi}_\ell)$.
4. Compute $\vec{x} \leftarrow \arg\min_{\vec{x}} \|\vec{v} - \Phi_t\vec{x}\|_2$.
5. Update the residual as $\vec{r} \leftarrow \vec{v} - \Phi_t\vec{x}$.
6. Increment t and return to step 2 if $t < m$.

(a) Explain step 2 geometrically. Why is it a reasonable heuristic for adding $\vec{\varphi}_\ell$ to Φ_{t-1}? [3 points]
NOTE: You can assume that $\|\vec{\varphi}_j\|_2 = 1$ for all $j \in \{1,\ldots,d\}$.

(b) Suppose between steps 3 and 4 we factor $\Phi_t = Q_tR_t$, where $Q_t \in \mathbb{R}^{n\times t}$ has orthonormal columns and $R_t \in \mathbb{R}^{t\times t}$ is upper-triangular. How can this factorization be used to carry out step 4? [2 points]

(c) Suppose we have the factorization $\Phi_{t-1} = Q_{t-1}R_{t-1}$ from the previous iteration. Propose a way to update to Q_t and R_t incrementally when $\vec{\varphi}_\ell$ gets added as a column to construct Φ_t in step 3. [5 points]

CIRCLE WHICH SIX OF PROBLEMS 2-7 YOU WANT GRADED ON PAGE 1.

Problem 6 (Integration and differentiation).

(a) Suppose $f(x) : [0,1] \to \mathbb{R}$ is evaluated on a system subject to numerical error. When might it be numerically stable to approximate $\int_0^1 f(x),\, dx$ but not $f'(0.5)$? [2 points]

(b) Consider the following formula for integrating with infinite bounds:

$$\int_0^\infty f(x)\,dx = \int_0^1 \frac{f(-\ln t)}{t}\,dt,$$

which allows for integration on a finite interval $[0,1]$ rather than an infinite interval $[0,\infty)$. When applying this reduction, would you integrate the right-hand side using closed or open quadrature? Why? [3 points]

(c) Suppose we wish to find a root of $f : \mathbb{R}^n \to \mathbb{R}^n$, but we do not know its Jacobian Df. Two ways we might find the root would be:

1. Newton's method, where we approximate Df using divided differences.
2. Broyden's method, which runs like Newton's method but updates a rough approximation J of Df using secants.

Compare the cost of a single iteration of these two techniques. If n is large and f is smooth but time-intensive to evaluate, which is likely preferable? [5 points]

CIRCLE WHICH SIX OF PROBLEMS 2-7 YOU WANT GRADED ON PAGE 1.

Problem 7 (Conjugate gradients and optimization). For all parts of this problem, assume line search finds a restricted minimizer in each step of optimization.

(a) Suppose $f(\vec{x}) : \mathbb{R}^n \to \mathbb{R}$ is smooth and bounded below. We can run gradient descent on f twice, starting from two different points $\vec{x}_0, \vec{x}_1 \in \mathbb{R}^n$. Will the two runs necessarily converge to the same point? Why? [2 points]

(b) Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite and $\vec{b} \in \mathbb{R}^n$. If $f(\vec{x})$ from part (a) satisfies $f(\vec{x}) = \vec{x}^\top A\vec{x} - 2\vec{x}^\top \vec{b}$, does your answer to part (a) change? Why? [2 points]

(c) Suppose we implement conjugate gradients to solve $A\vec{x} = \vec{b}$ for some symmetric positive definite $A \in \mathbb{R}^{n \times n}$ and $\vec{b} \in \mathbb{R}^n$. Unfortunately, A is poorly conditioned, and due to numerical error the search directions are no longer exactly A-conjugate. Does conjugate gradients necessarily converge in $\leq n$ steps in this case? Why? [2 points]

(d) Suppose we run the error-prone system from part (c). Conjugate gradients converges within numerical precision to a point $\vec{x}_0 \in \mathbb{R}^n$, but $\|A\vec{x}_0 - \vec{b}\|_2$ is relatively large. Hypothesize what went wrong, and propose a method for fixing this failure mode. [4 points]

THIS EXTRA CREDIT PROBLEM IS COMPLETELY OPTIONAL.

Extra credit. Suppose $\Omega \subseteq \mathbb{R}^2$ is closed and compact.

(a) Recall the Dirichlet energy defined in lecture:

$$E[u] \equiv \int_\Omega \|\nabla u(x,y)\|_2^2 \, dx \, dy.$$

We showed informally that Laplace's equation $v_{xx} + v_{yy} = 0$ minimizes $E[v]$ with respect to $v : \Omega \to \mathbb{R}$. Use a similar argument to relate the heat equation $u_t = u_{xx} + u_{yy}$ to gradient descent on $E[u]$. Assume Dirichlet boundary conditions $u|_{\partial\Omega} \equiv 0 \, \forall t \geq 0$ and that $u(x,y,t)$ is given for $(x,y,t) \in \Omega \times \{0\}$. [5 points]

(b) Consider the *inverse heat equation* $-u_t = u_{xx} + u_{yy}$.

(i) Is the inverse heat equation parabolic? [1 point]

☐ Yes ☐ No

(ii) Is the inverse heat equation well-posed? [1 point]

☐ Yes ☐ No

(iii) Propose an application of numerical methods for the inverse heat equation. [3 points]

THIS PAGE IS FOR ADDITIONAL WORK.

Problem number:

Problem number:

THIS PAGE IS FOR ADDITIONAL WORK.

Problem number:

Problem number: