Homework 4: SVD and Root-Finding

CS 205A: Mathematical Methods for Robotics, Vision, and Graphics (Fall 2013), Stanford University

Due Monday, November 4, at midnight

Problem 1 (30 points). *As promised in class:*

(a) *We'll start with some review from lecture. Derive the Broyden step by solving the following minimization in closed form for J_k:*

$$\begin{aligned} \textit{minimize}_{J_k} & \ \|J_k - J_{k-1}\|_{Fro}^2 \\ \textit{such that} & \ J_k(\vec{x}_k - \vec{x}_{k-1}) = f(\vec{x}_k) - f(\vec{x}_{k-1}) \end{aligned}$$

Hint: This problem was solved in class and is solved in the lecture notes.

(b) *Verify the Sherman-Morrison formula:*

$$(A + \vec{u}\vec{v}^\top)^{-1} = A^{-1} - \frac{A^{-1}\vec{u}\vec{v}^\top A^{-1}}{1 + \vec{v}^\top A^{-1}\vec{u}}$$

(c) *Assuming J_0 is the identity matrix, give a version of Broyden's root-finding algorithm that does not require matrix inversion.*

Problem 2 (40 points). *Take $\sigma_i(A)$ to be the i-th singular value of the square matrix $A \in \mathbb{R}^{n\times n}$. Define the* nuclear norm *of A to be*

$$\|A\|_* \equiv \sum_{i=1}^{n} \sigma_i(A).$$

Note: What follows is a tricky problem. Get started early, and feel free to ask questions on Piazza or in office hours. Also recall our mantra from lecture: "If a linear algebra problem is hard, substitute the SVD."

(a) *Show $\|A\|_* = tr(\sqrt{A^\top A})$, where trace of a matrix $tr(A)$ is the sum $\sum_i a_{ii}$ of its diagonal elements. For this problem, we will define the square root of a symmetric, positive semidefinite matrix M to be $\sqrt{M} \equiv XD^{1/2}X^\top$, where $D^{1/2}$ is the diagonal matrix containing (nonnegative) square roots of the eigenvalues of M and X contains the eigenvectors of $M = XDX^\top$.*

Hint (to get started): Write $A = U\Sigma V^\top$ and argue $\Sigma^\top = \Sigma$ in this case.

(b) *If $A, B \in \mathbb{R}^{n\times n}$, show $tr(AB) = tr(BA)$.*

(c) *Show $\|A\|_* = \max_{C^\top C = I} tr(AC)$. [Hint: Substitute the SVD of A and apply part (b).]*

(d) *Show that $\|A + B\|_* \le \|A\|_* + \|B\|_*$. [Hint: Use part (c).]*

(e) *Recall from lecture that minimizing* $\|A\vec{x} - \vec{b}\|_2^2 + \|\vec{x}\|_1$ *provides an alternative to Tikhonov regularization that can yield* sparse *vectors* $\vec{x}$ *under certain conditions. Assuming this is the case, in a few sentences explain informally why minimizing* $\|A - A_0\|_{Fro}^2 + \|A\|_*$ *over* A *for a fixed* $A_0 \in \mathbb{R}^{n\times n}$ *might yield a* low-rank *approximation of* A_0.

EC. *Read about the "low-rank matrix completion" problem and explain one practical application of its solution; the strategy in part (e) is one optimization approach to solving this problem.*

Problem 3 (15 points)**.** *Some shorter exercises:*

(a) *[Heath 5.11] Suppose you are using the secant method to find a root* x^* *of a nonlinear equation* $f(x) = 0$. *Show that if any iteration it happens to be the case that either* $x_k = x^*$ *or* $x_{k-1} = x^*$ *(but not both), then it will also be true that* $x_{k+1} = x^*$.

(b) *Show that adding a row to a matrix cannot decrease its largest or smallest singular value.*

Problem 4 (15 points)**.** *In this problem, we will derive a technique is known as Newton-Raphson division. Thanks to its fast convergence, it is often implemented in hardware for IEEE-754 floating point arithmetic.*

(a) *Show how the reciprocal* $\frac{1}{a}$ *of* $a \in \mathbb{R}$ *can be computed iteratively using Newton's method. Write your iterative formula in a way that requires at most two multiplications, one addition or subtraction, and no divisions.*

(b) *Take* x_k *to be the estimate of* $\frac{1}{a}$ *during the k-th iteration of Newton's method. If we define* $\varepsilon_k \equiv ax_k - 1$, *show that* $\varepsilon_{k+1} = -\varepsilon_k^2$.

(c) *Approximately how many iterations of Newton's method are needed to compute* $\frac{1}{a}$ *within d binary decimal points? Write your answer in terms of* ε_0 *and* d, *and assume* $|\varepsilon_0| < 1$.