Solution Midterm 1 Spring 2016

Please write your answers on these sheets, use the back sides if needed. Show your work. You can use a fact from the slides/book without having to prove it unless you are specifically asked to do so. Be organized and use readable handwriting. There is a page for scratch work at the end.

Solution 1 (Solution of optimization problems.) Give specific examples of functions $f_0 : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R}^n \to \mathbb{R}$ such that the optimization problem $\min_x f_0(x)$ subject to $f(x) \leq 0$ has the following properties. Only give one example per case for a total of three examples. Please no drawings. Give the formulae for f_0 and f.

(a) (5 pts.) The set of optimal solutions contains one point.

$n = 1,\ f_0(x) = x^2,\ f(x) = x.$

(b) (5 pts.) The set of optimal solutions contains an infinite number of points.

$n = 1,\ f_0(x) = 0,\ f(x) = x.$

(c) (5 pts.) The set of optimal solutions is empty and there is a constant $a \in \mathbb{R}$ such that $f_0(x) \geq a$ for all $x \in \mathbb{R}^n$.

$n = 1,\ f_0(x) = e^x,\ f(x) = x,\ a = 0.$

Solution 2 (Matrix norms.) (15 pts.) A matrix $A \in \mathbb{R}^{m,n}$ with rank r has singular values $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$. Prove that the spectral norm satisfies $\|A\|_2^2 = \sigma_1^2$.

Theorem 4.3 states that

$$\frac{x^\top A^\top A x}{x^\top x} \leq \lambda_1(A^\top A), \text{ for all } x \neq 0,$$

where $\lambda_1(A^\top A)$ is the largest eigenvalue of $A^\top A$. Moreover, for $x = u_1$, with u_1 a unit-norm eigenvalue of $A^\top A$ corresponding to $\lambda_1(A^\top A)$, we have that the inequality holds with equality. That is,

$$\frac{\|Au_1\|_2^2}{\|u_1\|_2^2} = \lambda_1(A^\top A).$$

Thus, $\|Au_1\|_2^2 = \lambda_1(A^\top A) = \sigma_1^2$, where σ_1 is the largest singular value of A by the SVD theorem.

Solution 3 (Matrix approximation.) For a given $A \in \mathbb{R}^{m,n}$, with $\text{rank}(A) = r$, consider the problem

$$\min_{A_k \in \mathbb{R}^{m,n}} \|A - A_k\|_F^2 \text{ subject to } \text{rank}(A_k) = k.$$

Let $\Sigma_{i=1}^r \sigma_i u_i v_i^\top$ be a singular value decomposition of A. For $k \leq r$, it is known that an optimal solution of the problem is $A_k = \Sigma_{i=1}^k \sigma_i u_i v_i^\top$.

(a) (5 pts.) Suppose that $r = 4$ and $\sigma_1 = 4$, $\sigma_2 = 2$, $\sigma_3 = 2$, and $\sigma_4 = 1$. Quantify the relative error in A_k compared to the "true" matrix A for $k = 1, 2, 3,$.

$$\frac{\|A - A_k\|_F^2}{\|A\|_F^2} = \frac{\sigma_{k+1}^2 + ... + \sigma_r^2}{\sigma_1^2 + ... + \sigma_r^2}.$$

For $k = 1$, this ratio becomes $(4 + 4 + 1)/(16 + 4 + 4 + 1) = 9/25$

For $k = 2$, this ratio becomes $(4 + 1)/(16 + 4 + 4 + 1) = 5/25$

For $k = 3$, this ratio becomes $1/(16 + 4 + 4 + 1) = 1/25$

(b) (10 pts.) Suppose $m \geq n$ and $\text{rank}(A) = n$. Formulate an optimization problem that determines how "far" A is from being of rank $n - 1$. Solve this problem and obtain an explicit expression for a matrix $B \in \mathbb{R}^{m,n}$ such that $A + B$ has rank $n - 1$. (Ignore what was given in part a.)

The problem becomes

$$\min_{B \in \mathbb{R}^{m,n}} \|B\|_F^2 \text{ subject to } \text{rank}(A + B) = n - 1.$$

If we set $k = n-1$ and $A_k = A + B$, then the new problem is of the same form as the new original problem and thus we have an optimal solution $A_k = \sum_{i=1}^{n-1} \sigma_i u_i v_i^\top$. An optimal $B = A_k - A = -\sigma_n u_n v_n^\top$.

Solution 4 (Optimization over norm balls.) (10 pts.) For a given $y \in \mathbb{R}^n$, derive an optimal solution of the problem $\max x^\top y$ subject to $\|x\|_\infty \leq 1$.

Since $x^\top y = \sum_{i=1}^n x_i y_i$ and $|x_i| \leq 1$ for all i, we select $x_i = 1$ when $y_i > 0$, $x_i = -1$ when $y_i < 0$, and x_i arbitrarily when $y_i = 0$ to achieve a maximum solution. The maximum value is then $\sum_{i=1}^n |y_i| = \|y\|_1$.

Solution 5 (Projection on a hyperplane.) Consider the hyperplane $\{z \in \mathbb{R}^n \; : \; a^\top z = b\}$, $a \neq 0$, and a point $y \in \mathbb{R}^n$.

(a) (10 pts.) Determine the Euclidean projection of y onto the hyperplane.

We need to have that the projection y^* satisfies $a^\top y^* = b$ and $(y - y^*)$ is perpendicular to the hyperplane, i.e., $y - y^* = \alpha a$ for some $\alpha \in \mathbb{R}$. Pre-multiplying the last condition with $a^\top$, we obtain that

$$a^\top (y - y^*) = \alpha \|a\|_2^2.$$

Substituting in $a^\top y^* = b$, this leads to $a^\top y - b = \alpha \|a\|_2^2$ and

$$\alpha = \frac{a^\top y - b}{\|a\|_2^2}.$$

The projection of y is therefore

$$y^* = y - \alpha a = \frac{a^\top y - b}{\|a\|_2^2} a.$$

(b) (5 pts.) Determine the Euclidean distance between y and its projection on the hyperplane.

Plugging in y^* from above, we find that

$$\|y - y^*\|_2 = |a| \|a\|_2 = \frac{|a^\top y - b|}{\|a\|_2}.$$

Solution 6 (Properties of dyad.) Let $x, y \in \mathbb{R}^n$, both not identical to the zero vector, and $A = xy^\top \in \mathbb{R}^{n,n}$.

(a) (5 pts.) Determine an eigenvalue and an eigenvector of A.

Eigenvalue $\lambda = y^\top x$ and eigenvector $u = x$ work because, $Au = xy^\top x = u\lambda$.

(b) (5 pts.) We know that A has rank one. Write a proof of this fact.

$\mathcal{R}(A) = \{z \in \mathbb{R}^n \ : \ z = Av, v \in \mathbb{R}^n\}$. Since $Av = xy^\top v = \gamma x$ for $\gamma = y^\top v$, the range of A is simply a line. Thus, there is only one linearly independent column in A.

(c) (5 pts.) What is the dimension of $\mathcal{N}(A)$?

The dimension of $\mathcal{N}(A) = n - \operatorname{rank} A = n - 1$ by the fundamental theorem of linear algebra.

(d) (5 pts.) Compute a singular value decomposition of A and write it in compact form.

Take $\sigma = \|x\|_2\|y\|_2$, $u = x/\|x\|_2$, and $v = y/\|y\|_2$. Clearly, $A = \sigma uv^\top$. Moreover, $u^\top u = 1$, $v^\top v = 1$, $Av = xy^\top y/\|y\|_2 = \sigma u$, and $u^\top A = x^\top xy^\top/\|x\|_2 = \sigma v$. Thus, σ, u, v is a SVD of A.

Solution 7 (Bound on a polynomial's derivative.) (10 pts.) For $w \in \mathbb{R}^{k+1}$, we define the polynomial p_w, with values

$$p_w(x) \doteq w_1 + w_2 x + \ldots + w_{k+1}x^k.$$

Prove that

$$\forall\, x \in [-1, 1] \ : \ \left|\frac{\mathrm{d}p_w(x)}{\mathrm{d}x}\right| \leq k^{3/2}\|v\|_2,$$

where $v = (w_2, \ldots, w_{k+1}) \in \mathbb{R}^k$.

With $z = (1, 2, \ldots, k)$, Cauchy-Schwartz inequality gives that

$$\begin{aligned}
\left|\frac{\mathrm{d}p_w(x)}{\mathrm{d}x}\right| &= \left|w_2 + 2w_3 x + \ldots + kw_{k+1}x^{k-1}\right| \\
&\leq |w_2| + 2|w_3| + \ldots + k|w_{k+1}| \\
&= |v^\top z| \\
&\leq \|v\|_2 \cdot \|z\|_2.
\end{aligned}$$

The conclusion follows after realizing that

$$\|z\|_2 = \sqrt{1 + 4 + \ldots + k^2} \leq \sqrt{k \cdot k^2} = k^{3/2}.$$