Midterm

Michael Nelson

Throughout this homework, $\|\cdot\|$ denotes the ℓ_2-norm (except in part 1 of problem 1 where we use the ℓ_∞ norm). If $v, w \in \mathbb{R}^n$, then we set $\langle v, w\rangle = v^\top w$. Also we set $\varepsilon = \varepsilon_{\text{mach}}$ to be the machine coefficient. The exam took me roughly 6 hours to finish.

Problem 1

Exercise 1. Solve the following.

1. Let $f(\boldsymbol{x}) = x_1^2 \ln x_2$ where $x_2 > 0$. Find the relative condition number of f. If $x_1 \approx 1$, for what values of x_2 is this evaluation ill-conditioned?

2. Let $A \in \mathbb{R}^{n\times n}$ and (λ, v) be an eigenpair such that $Av = \lambda v$. Let $(\widehat{\lambda}, \widehat{v})$ be a numerically computed approximation to (λ, v). Assume that $\|v\| = \|\widehat{v}\| = 1$. Find the vector x such that $(\widehat{\lambda}, \widehat{v})$ is an eigenpair of $A + \Delta A$, where $\Delta A = x\widehat{v}^H$. Give the expression of $\|\Delta A\|$ (should not contain x), and explain why $\|A\widehat{v} - \widehat{\lambda}\widehat{v}\| \leq O(\|A\|\varepsilon)$ means $(\widehat{\lambda}, \widehat{v})$ is computed by a backward stable algorithm.

Solution 1. 1. In this part of the problem, we use the ℓ_∞ norm $\|\cdot\|_\infty = \|\cdot\|$. Since f is differentiable we see that

$$
\begin{aligned}
\kappa_f(\boldsymbol{x}) &= \frac{\|\mathrm{J}_f(\boldsymbol{x})\|\|\boldsymbol{x}\|}{|f(\boldsymbol{x})|} \\
&= \frac{\left\|\begin{pmatrix}2x_1\ln x_2 & x_1^2/x_2\end{pmatrix}\right\| \|\boldsymbol{x}\|}{|x_1^2 \ln x_2|} \\
&= \frac{\left(|2x_1 \ln x_2| + |x_1^2/x_2|\right)\max\{|x_1|, |x_2|\}}{|x_1^2 \ln x_2|} \\
&= \frac{\left(|2\ln x_2| + |x_1/x_2|\right)\max\{|x_1|, |x_2|\}}{|x_1 \ln x_2|}, \\
&= \frac{2\max\{|x_1|, |x_2|\}}{|x_1|} + \frac{\max\{|x_1|, |x_2|\}}{|x_2 \ln x_2|},
\end{aligned}
$$

Thus we have

$$
\kappa_f(\boldsymbol{x}) = \begin{cases} \frac{2x_2}{|x_1|} + \frac{1}{\ln x_2} & \text{if } x_2 \geq |x_1| \\ 2 + \frac{|x_1|}{x_2|\ln x_2|} & \text{if } |x_1| \geq x_2 \end{cases}
$$

Now we assume that $x_1 \approx 1$. Then

$$
\kappa_f(\boldsymbol{x}) \approx \begin{cases} 2x_2 + \frac{1}{\ln x_2} & \text{if } x_2 \geq x_1 \approx 1 \\ 2 + \frac{1}{x_2|\ln x_2|} & \text{if } x_2 \leq x_1 \approx 1 \end{cases}
$$

Let us analyze for which x_2 is this problem is ill-conditioned. First we consider the case where $x_2 \geq x_1 \approx 1$. Then the problem is ill-conditioned if and only if

$$
\kappa_f(\boldsymbol{x}) = 2x_2 + \frac{1}{\ln x_2}
$$

is large, which happens if and only if x_2 is large. Next we consider the case where $0 < x_2 \leq x_1 \approx 1$. Then the problem is ill-conditioned if and only if

$$
2 + \frac{1}{x_2\,|\ln x_2|}
$$

is large, which happens if and only if $\frac{1}{x_2|\ln x_2|}$ is large which happens if $x_2 \approx 0$ or $x_2 \approx 1$.

2. We have

$$\begin{aligned}(A+\Delta A)\widehat{v} = \widehat{\lambda}\widehat{v} &\iff (A + x\widehat{v}^\top)\widehat{v} = \widehat{\lambda}\widehat{v} \\ &\iff A\widehat{v} + x\|\widehat{v}\|^2 = \widehat{\lambda}\widehat{v} \\ &\iff A\widehat{v} + x = \widehat{\lambda}\widehat{v} \\ &\iff x = \widehat{\lambda}\widehat{v} - A\widehat{v}.\end{aligned}$$

Thus we have

$$\begin{aligned}\|\Delta A\| &= \|x\|\|\widehat{v}\| \\ &= \|x\| \\ &= \|\widehat{\lambda}\widehat{v} - A\widehat{v}\|.\end{aligned}$$

In particular, $\|\widehat{\lambda}\widehat{v} - A\widehat{v}\| = O(\|A\|\varepsilon)$ means

$$\frac{\|\Delta A\|}{\|A\|} = \frac{\|\widehat{\lambda}\widehat{v} - A\widehat{v}\|}{\|A\|} = O(\varepsilon),$$

which shows that $(\widehat{\lambda}, \widehat{v})$ is computed by a backward stable algorithm.

1 Problem 2

Exercise 2. Solve the following:

1. Assume that $U, V \in \mathbb{R}^{n\times p}$ $(p < n)$ have full rank, and $V^\top U$ is nonsingular. Show that $P = 1 - U(V^\top U)^{-1}V^\top$ is a projector, and find the range and null space of P. What can we say about P and $\|P\|$ if $\operatorname{range}(U) = \operatorname{range}(V)$?
2. Let $x \in \mathbb{R}^n$ and consider the vector $z = (0_{n-1}, \|x\|, x)^\top \in \mathbb{R}^{2n}$. Find the Householder reflector $H = 1 - 2vv^\top$ that reduces z such that Hz is a multiple of e_1 (sufficient to find the expression of v). For $y = (0_n, x)^\top \in \mathbb{R}^{2n}$, give the simplified expression of Hy.
3. Given a 6×4 matrix A with all nonzero entries, illustrate the procedure of Golub-Kahan bidiagonalization, and explain how to compute all singular values of A.

Solution 2. 1. Observe that

$$\begin{aligned}P^2 &= (1 - U(V^\top U)^{-1}V^\top)(1 - U(V^\top U)^{-1}V^\top) \\ &= 1 - 2U(V^\top U)^{-1}V^\top + (U(V^\top U)^{-1}V^\top)(U(V^\top U)^{-1}V^\top) \\ &= 1 - 2U(V^\top U)^{-1}V^\top + U(V^\top U)^{-1}(V^\top U)(V^\top U)^{-1}V^\top \\ &= 1 - 2U(V^\top U)^{-1}V^\top + U(V^\top U)^{-1}V^\top \\ &= 1 - U(V^\top U)^{-1}V^\top \\ &= P.\end{aligned}$$

It follows that P is a projector. We have

$$\begin{aligned}x \in \operatorname{range}(P) &\iff Px = x && \text{since } P \text{ is a projector} \\ &\iff x - U(V^\top U)^{-1}V^\top x = x \\ &\iff U(V^\top U)^{-1}V^\top x = 0 \\ &\iff (V^\top U)^{-1}V^\top x = 0 && \text{since } U\colon \mathbb{R}^p \to \mathbb{R}^n \text{ is injective} \\ &\iff V^\top x = 0 && \text{since } (V^\top U)^{-1}\colon \mathbb{R}^p \to \mathbb{R}^p \text{ is injective} \\ &\iff x \in \operatorname{null}(V^\top).\end{aligned}$$

Thus we have $\text{range}(P) = \text{null}(V^\top)$. Similarly we claim that $\text{null}(P) = \text{range}(U)$. To see this, first suppose $x \in \text{range}(U)$, so $x = Uy$. Then

$$\begin{aligned}
Px &= x - U(V^\top U)^{-1}V^\top x \\
&= Uy - U(V^\top U)^{-1}V^\top x \\
&= U(y - (V^\top U)^{-1}V^\top x) \\
&= U(y - (V^\top U)^{-1}V^\top Uy) \\
&= U(y - y) \\
&= 0
\end{aligned}$$

implies $x \in \text{null}(P)$. Thus $\text{range}(U) \subseteq \text{null}(P)$. Conversely, suppose that $x \in \text{null}(P)$. Then

$$\begin{aligned}
0 &= Px \\
&= x - U(V^\top U)^{-1}V^\top x \\
&= x - Uy
\end{aligned}$$

where $y = (V^\top U)^{-1}V^\top x$ implies $x \in \text{range}(U)$. Thus $\text{null}(P) \subseteq \text{range}(U)$. Thus we have shown

$$\text{range}(P) = \text{null}(V^\top) \quad \text{and} \quad \text{null}(P) = \text{range}(U).$$

We now assume that $\text{range}(U) = \text{range}(V)$. In this case, one has the decomposition

$$\mathbb{R}^n \cong \text{null}(P) \oplus \text{range}(P) = \text{range}(V) \oplus \text{null}(V^\top). \tag{1.1}$$

Indeed, suppose $x \in \text{range}(V) \cap \text{null}(V^\top)$. Choose y such that $Vy = x$. Then note that

$$\begin{aligned}
\|x\|^2 &= \|Vy\|^2 \\
&= (Vy)^\top Vy \\
&= y^\top V^\top Vy \\
&= y^\top V^\top x \\
&= 0
\end{aligned}$$

implies $x = 0$. Thus $\text{range}(V) \cap \text{null}(V^\top) = 0$. Furthermore, note that

$$\dim(\text{range}(V)) = p \quad \text{and} \quad \dim(\text{null}(V^\top)) = n - p$$

since V has full rank. Thus we obtain the decomposition (1.1) as claimed. So every $x \in \mathbb{R}^n$ can be expressed uniquely as $x = Vy + z$ where $y \in \mathbb{R}^p$ and where $z \in \text{null}(V^\top)$. In fact, we claim (1.1) is an orthogonal decomposition. Indeed, suppose $Vy \in \text{range}(V)$ and $z \in \text{null}(V^\top)$. Then note that

$$\langle Vy, z \rangle = (Vy)^\top z = y^\top V^\top z = 0.$$

Thus (1.1) is an orthogonal decomposition and P is the orthogonal projection map onto $\text{null}(V^\top)$. In this case, we have $\|P\| = 1$. Indeed,

$$\begin{aligned}
\|Px\| &= \|P^2 x\| \\
&\leq \|P\|\|Px\|
\end{aligned}$$

shows $\|P\| \geq 1$. Conversely, by the Pythagorean theorem we have

$$\begin{aligned}
\|x\|^2 &= \|Px\|^2 + \|x - Px\|^2 \\
&\geq \|Px\|^2,
\end{aligned}$$

which implies $\|Px\| \leq x$ which implies $\|P\| \leq 1$.

2. Let $\widehat{z} = z/\|z\|$ and let $v = \widehat{z} - e_1$. Note that

$$\begin{aligned}\langle v, \widehat{z}\rangle &= \langle \widehat{z} - e_1, \widehat{z}\rangle \\ &= \langle \widehat{z}, \widehat{z}\rangle - \langle e_1, \widehat{z}\rangle \\ &= 1 - 0 \\ &= 1.\end{aligned}$$

Similarly note that since $\langle \widehat{z}, e_1\rangle = 0$, we have

$$\begin{aligned}\|v\|^2 &= \|\widehat{z}\|^2 + \|e_1\|^2 \\ &= 1 + 1 \\ &= 2,\end{aligned}$$

Therefore we have

$$\begin{aligned}H_v(z) &= \|z\| H_v(\widehat{z}) \\ &= \|z\| \left(\widehat{z} - 2\frac{\langle v, \widehat{z}\rangle}{\|v\|^2} v\right) \\ &= \|z\| \left(\widehat{z} - v\right) \\ &= \|z\|(\widehat{z} - (\widehat{z} - e_1)) \\ &= \|z\| e_1 \\ &= \sqrt{2}\|x\| e_1, && \text{since } \|z\|^2 = 2\|x\|^2.\end{aligned}$$

In other words, $H = H_v = 1 - 2\widehat{v}\widehat{v}^\top$ where $\widehat{v} = v/||v\|$ is the Hausdorff reflector which reduces z such that

$$Hz = \sqrt{2}\|x\| e_1.$$

Now suppose that $y = (0_n, x)^\top \in \mathbb{R}^{2n}$, then we have

$$\begin{aligned}\langle v, y\rangle &= \langle \widehat{z} - e_1, y\rangle \\ &= \langle \widehat{z}, y\rangle - \langle e_1, y\rangle && \text{since } \langle e_1, y\rangle = 0 \\ &= \langle \widehat{z}, y\rangle \\ &= \frac{1}{\|z\|}\langle z, y\rangle \\ &= \frac{1}{\sqrt{2}\|x\|}\|x\|^2 && \text{since } \langle z, y\rangle = \|x\|^2 \\ &= \frac{1}{\sqrt{2}}\|x\|,\end{aligned}$$

Thus we have

$$\begin{aligned}H_v(y) &= y - 2\frac{\langle v, y\rangle}{\|v\|^2} v \\ &= y - \langle v, y\rangle v \\ &= y - \frac{1}{\sqrt{2}}\|x\| \left(\frac{z}{\|z\|} - e_1\right) \\ &= y - \frac{1}{\sqrt{2}}\|x\| \left(\frac{z}{\sqrt{2}\|x\|} - e_1\right) \\ &= y - \frac{z}{2} + \frac{1}{\sqrt{2}}\|x\| e_1 \\ &= \begin{pmatrix} \|x\|/\sqrt{2} \\ 0_{n-2} \\ -\|x\|/2 \\ x/2 \end{pmatrix}.\end{aligned}$$

3. Let A be a 6×4 matrix with nonzero entries:

$$A = \begin{pmatrix} \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{pmatrix},$$

where the $\star$'s indicate nonzero entries. The Golub-Kahan bidiagonalization procedure is illustrated as follows:

$$\begin{aligned}
A = \begin{pmatrix} \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{pmatrix} &\to \begin{pmatrix} \star & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \end{pmatrix} = U_1 A \\
&\to \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \end{pmatrix} = U_1 A V_1^\top \\
&\to \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \end{pmatrix} = U_2 U_1 A V_1^\top \\
&\to \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & \star & 0 \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & \star & \star \end{pmatrix} = U_2 U_1 A V_1^\top V_2^\top \\
&\to \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & \star & 0 \\ 0 & 0 & \star & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & \star \end{pmatrix} = U_3 U_2 U_1 A V_1^\top V_2^\top \\
&\to \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & \star & 0 \\ 0 & 0 & \star & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = U_4 U_3 U_2 U_1 A V_1^\top V_2^\top \\
&= \begin{pmatrix} B \\ 0 \end{pmatrix}
\end{aligned}$$

where the U_i and V_j are appropriately chosen Householder transformations. Thus we obtain

$$\begin{pmatrix} B \\ 0 \end{pmatrix} = UAV^\top,$$

where B is a bidiagonal matrix:

$$B = \begin{pmatrix} b_1 & c_1 & 0 & 0 \\ 0 & b_2 & c_2 & 0 \\ 0 & 0 & b_3 & c_3 \\ 0 & 0 & 0 & b_4 \end{pmatrix}.$$

Thus to obtain the singular values of A, we just need to calculate the eigenvalues of

$$B^\top B = \begin{pmatrix} b_1^2 & b_1c_1 & 0 & 0 \\ b_1c_1 & b_2^2 + c_1^2 & b_2c_2 & 0 \\ 0 & b_2c_2 & b_3^2 + c_2^2 & b_3c_3 \\ 0 & 0 & b_3c_3 & b_4^2 + c_3^3 \end{pmatrix},$$

and the singular values of A will be the square root of the eigenvalues of $B^\top B$.

2 Problem 3

Exercise 3. Solve the following:

1. Let $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, let $A^\dagger = (A^\top A)^{-1}A^\top$. Show that $\|A^\dagger\| = 1/\sigma_n(A)$ (assume that A has full column rank).

2. Let $A = \begin{pmatrix} A_1 & A_2 \end{pmatrix}^\top$ where $A_1 \in \mathbb{R}^{n\times n}$ is nonsingular. Show that $\sigma_n(A) \geq \sigma_n(A_1)$ (explore the relation between $\|Ax\|/\|x\|$ and $\|A_1x\|/\|x\|$) and $\|A^\dagger\| \leq \|A_1^{-1}\|$.

3. Define the numerial rank of $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ as $\operatorname{rank}(A, \varepsilon) = \max\{k \mid \sigma_k \geq \varepsilon\}$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n)$. If A has numerical rank $k < n$ for a given ε, find a numerically full rank B satisfying $\inf_{\operatorname{rank}(B,\varepsilon)=n} \|A - B\|_F$ and show that $\|B - A\|_F \leq \varepsilon\sqrt{n-k}$.

Solution 3. 1. Let $A = U\Sigma V^\top$ be an SVD of A where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n, 0_{m-n})$ where $\sigma_1 \geq \cdots \geq \sigma_n$. Then

$$\begin{aligned} A^\dagger &= (A^\top A)^{-1}A^\top \\ &= (V\Sigma^\top U^\top U\Sigma V^\top)^{-1}(V\Sigma^\top U^\top) \\ &= (V\Sigma^\top \Sigma V^\top)^{-1}(V\Sigma^\top U^\top) \\ &= (\Sigma^\top \Sigma V V^\top)^{-1}(V\Sigma^\top U^\top) \\ &= (\Sigma^\top \Sigma)^{-1}(V\Sigma^\top U^\top) \\ &= V((\Sigma^\top \Sigma)^{-1}\Sigma^\top)U^\top \\ &= V\Sigma^\dagger U^\top, \end{aligned}$$

shows that $A^\dagger = V\Sigma^\dagger U^\top$ is an SVD of $A^\dagger$, where we used the fact that $\Sigma^\top\Sigma$ is a diagonal matrix (and thus commutes with all other matrices) and where we set $\Sigma^\dagger = (\Sigma^\top\Sigma)^{-1}\Sigma^\top$. A straightforward computation shows that $\Sigma^\dagger$ has the form:

$$\Sigma^\dagger = \begin{pmatrix} \operatorname{diag}(\sigma_1^{-1}, \ldots, \sigma_n^{-1}) & 0_{m-n} \end{pmatrix}.$$

Thus the singular values of $A^\dagger$ are $\sigma_1^{-1}, \ldots, \sigma_n^{-1}$. In particular, we have

$$\|A^\dagger\| = \max\{\sigma_1^{-1}, \ldots, \sigma_n^{-1}\} = \sigma_n^{-1}.$$

2. Observe that

$$\|Ax\|^2 = \|A_1x\|^2 + \|A_2x\|^2 \geq \|A_1x\|^2$$

implies $\|Ax\| \geq \|A_1x\|$ for all x. Thus

$$\begin{aligned} \sigma_n(A) &= \min\{\|Ax\|/\|x\| \mid x \neq 0\} \\ &\geq \min\{\|A_1x\|/\|x\| \mid x \neq 0\} \\ &= \sigma_n(A_1). \end{aligned}$$

In particular, this implies

$$\|A^\dagger\| = \sigma_n(A)^{-1} \leq \sigma_n(A_1)^{-1} = \|A_1^{-1}\|$$

by part 1 of this problem.

3. Let $A = U\Sigma V^\top$ be an SVD of A and set $B = U\widetilde{\Sigma}V^\top$ where $\widetilde{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_k, \varepsilon, \ldots, \varepsilon)$. Then note that

$$B - A = U(\widetilde{\Sigma} - \Sigma)V^\top$$

is an SVD of $B - A$ where $\widetilde{\Sigma} - \Sigma = \operatorname{diag}(0_k, \varepsilon - \sigma_{k+1}, \ldots, \varepsilon - \sigma_n)$. In particular, note that

$$\begin{aligned}
\|B - A\|_{\mathrm{F}} &= \sqrt{(\varepsilon - \sigma_{k+1})^2 + \cdots + (\varepsilon - \sigma_n)^2} \\
&\leq \sqrt{\varepsilon^2 + \cdots + \varepsilon^2} \\
&= (\sqrt{n-k})\varepsilon.
\end{aligned}$$