Idil_sahin_math56_midterm

5:50 PM

Math 56 Winter 2025 Midterm Exam

Student name (print): IDIL SAHIN

Date: 02/23/2025

Start/end time: 5:50pm - 11:11:pm

This exam contains 16 pages (including this cover page) and 8 questions. The total number of possible points is 63.

- **Take the exam in one six-hour sitting,** and mark your starting/ending times at the top of the exam in the space provided. The exam should not take the full six hours. You are allowed to take reasonably short breaks.
- **The exam is open-book and open-note, but NOT "open-internet" or "open-neighbor".** Discussion about the exam with other students (or other people) is not allowed. This includes chatbots such as ChatGPT. An exception to the internet rule are the YouTube videos I have prepared for the course, as well as the classroom lecture recordings on Panapto.
- **Organize your work**, in a reasonably neat and coherent way, in the space provided. Work scattered all over the page without a clear ordering will receive less credit. *There may be more blank space provided than is actually needed to answer each question.*
- **Use of calculators/code is permitted.**
- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations may still receive partial credit.
- **Provide exact answers** unless otherwise instructed.
- **Simplify all answers as much as possible.**
- **Clearly identify your final answer for each problem from any scratch work.**

Do not write in the table to the right. **Good luck!**

Question	Points	Score
1	10	
2	12	
3	5	
4	5	
5	5	
6	8	
7	12	
8	6	
Total:	63	

1. (10 points) Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Prove that

$$\mathbf{A}^{\dagger} = \lim_{\delta\to 0^+} \left(\mathbf{A}^T\mathbf{A} + \delta\mathbf{I}\right)^{-1}\mathbf{A}^T. \tag{1}$$

Hint: use the SVD, and pick a convenient matrix norm.

Consider the Frobenius norm

$\lim_{\delta\to 0} \| (A^T A + \delta I)^{-1} A^T - A^{\dagger} \|_F = e \in \mathbb{R}$, subsequently if $e = 0$, the matrices are equal.

Let $A \in \mathbb{R}^{m\times n}$ have the SVD decomposition, where;

$$A = U \Sigma V^T$$

- $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices
- $\Sigma \in \mathbb{R}^{m\times n}$ is a diagonal matrix with singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_r > 0$, where $r = \text{rank}(A)$ and zero values everywhere else.

Now observe that:

$$A^T = (V \Sigma^T U^T)$$

which gives us $A^T A = V \Sigma^T U^T (U \Sigma V^T)$

as U, U^T are orthogonal $= V \Sigma^T \Sigma V^T$.

Now recall that $VI = V$, $VIV^T = I$, as V is orthogonal.

plugging this back gives us:

$$A^T A + \delta I = V \Sigma^T \Sigma V^T + V \delta I V^T$$
$$= V(\Sigma^T \Sigma + \delta I) V^T$$

now consider:

$$(A^T A + \delta I)^{-1} = (V(\Sigma^T \Sigma + \delta I) V^T)^{-1}$$ V is orthogonal

$$= V(\Sigma^T \Sigma + \delta I) V^T$$

$$(A^TA + \delta I)^{-1}A^T = V(\Sigma^T\Sigma + \delta I)^{-1} V^T (V^T \Sigma^T U^T)$$
$$= V(\Sigma^T\Sigma + \delta I)^{-1} \Sigma^T U^T$$

Recall that $A^+ = V\Sigma^+ U^T$ subsequently focusing on Σ^+ and $(\Sigma^T\Sigma + \delta I)^{-1}\Sigma^T$ suffices, as V, V^T do not affect singular values and the F-norm.

Observe that:

$$\Sigma^T\Sigma + \delta I = \begin{bmatrix} \sigma_1^2+\delta & & & & & \\ & \sigma_2^2+\delta & & & & \\ & & \ddots & & & \\ & & & \sigma_r^2+\delta & & \\ & & & & \delta & \\ & & & & & \ddots \;\delta \end{bmatrix}$$

subsequently its inverse must be:

$$(\Sigma^T\Sigma + \delta I)^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2+\delta} & & & \\ & \ddots & & \\ & & \frac{1}{\delta} & \\ & & & \frac{1}{\delta} \end{bmatrix}$$

σ_1

and
$$(\Sigma^T\Sigma + \delta I)^{-1}\Sigma^T = \begin{bmatrix} \frac{\sigma_1}{\sigma_1 + \delta} & & & \\ & \ddots \; \frac{\sigma_r}{\sigma_r + \delta} & & \\ & & 0 & \\ & & & 0 \end{bmatrix}$$

now recall that
$$\Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots \; \frac{1}{\sigma_r} & & \\ & & 0 & \\ & & & 0 \end{bmatrix}:$$

now for every non-zero diagonal entry σ_i, we have

$$\lim_{\delta \to 0} \frac{\sigma_i}{\sigma_i^2+\delta} - \frac{1}{\sigma_i} \rightarrow \frac{\sigma_i}{\sigma_i^2} - \frac{1}{\sigma_i} = 0$$

as all eigenvalues are 0, $\sum_{i=1}^{r}\left(\frac{\sigma_i}{\sigma_i^2+\delta} - \frac{1}{\sigma_i}\right) = 0 = e$

which proves our statement.

2. (12 points) *Part (a):* Find an elementary lower triangular matrix $\mathbf{M}$ such that

$$\mathbf{M}\begin{bmatrix}2\\2\\1\end{bmatrix}=\begin{bmatrix}2\\0\\0\end{bmatrix}. \tag{2}$$

Part (b): Find the Householder reflector $\mathbf{H}=\mathbf{I}-2\mathbf{u}\mathbf{u}^T\in\mathbb{R}^{3\times3}$ (with $\mathbf{u}$ a unit vector) such that

$$\mathbf{H}\underset{x}{\begin{bmatrix}2\\2\\1\end{bmatrix}}=\underset{x'}{\begin{bmatrix}\star\\0\\0\end{bmatrix}}, \tag{3}$$

where "$\star$" denotes a generic nonzero entry. You can simply state the vector $\mathbf{u}$ and do not need to explicitly compute the entries of $\mathbf{H}$.

Part (c): Find (you do not need to give the entries explicitly) an orthogonal matrix $\mathbf{Q}\in\mathbb{R}^{5\times5}$ such that

$$\mathbf{Q}\overset{v}{\begin{bmatrix}5\\4\\2\\2\\1\end{bmatrix}}=\overset{v'}{\begin{bmatrix}\star\\\star\\\star\\0\\0\end{bmatrix}}\begin{matrix}x\\y\\z\\ \\ \end{matrix} \tag{4}$$

a) Recall that an elementary lower triangular matrix is defined as:

$$M=\begin{bmatrix}1&0&0\\-l_{31}&1&0\\-l_{23}&0&1\end{bmatrix}$$

now observe that:

$$\begin{bmatrix}1&0&0\\-l_{31}&1&0\\-l_{23}&0&1\end{bmatrix}\begin{bmatrix}2\\2\\1\end{bmatrix}=\begin{bmatrix}2\\-2l_{21}+2\\-2l_{23}+1\end{bmatrix}=\begin{bmatrix}2\\0\\0\end{bmatrix}$$

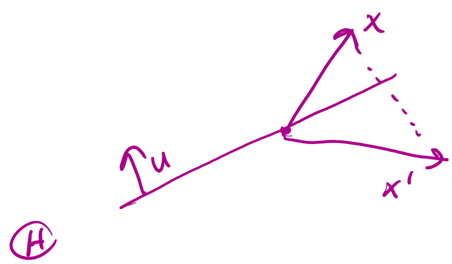
This gives us:

$$-2l_{21}=-2 \rightarrow -l_{21}=-1$$

$$-2l_{23}=-1 \rightarrow l_{23}=-\tfrac{1}{2}$$

Subsequently our matrix is: $\begin{bmatrix}1&0&0\\-1&1&0\\-1/2&0&1\end{bmatrix}$

b) Recall that a Householder reflector $H = I - 2uu^T$ reflects any vector x through the plane defined by u (perpendicular to u)



For ease of computation, let $\|x'\| = \|x\|$, which gives us $\|x'\| = \sqrt{2^2+2^2+1} = 3$, thus $x' = \langle 3,0,0 \rangle$.

Now notice that the difference of these vectors x, x' gives us a vector parallel to u.

Subsequently we get:

$$u = \frac{x - x'}{\|x - x'\|} = \frac{\langle 2,2,1 \rangle - \langle 3,0,0 \rangle}{\sqrt{(-1)^2 + 2^2 + 1^2}} = \frac{1}{\sqrt{6}} \langle -1, 2, 1 \rangle$$

c) using the same reasoning as part b), now observe we can choose $x^2+y^2+z^2 = 5^2+4^2+\underbrace{2^2+2^2+1}$

$x^2 = 5^2$ $\quad 4^2 = y^2$ $\quad 3^2 = z^2$

This gives us $v' = \langle 5, 4, 3, 0, 0 \rangle$

Deriving u, we get:

$$u = \frac{v - v'}{\| v - v' \|} = \frac{\langle 5,4,2,2,1 \rangle - \langle 5,4,3,0,0 \rangle}{\sqrt{(-1)^2 + 2^2 + 1^2}}$$

$$= \frac{1}{\sqrt{6}} \langle 0,0,-1,2,1 \rangle$$

$H = I - 2uu^T$,

3. (5 points) Let $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ be the economic QR decomposition where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. Explain how the Cholesky factorization $\mathbf{A} = \mathbf{L}\mathbf{L}^T$ can be obtained directly from the QR decomposition with no additional cost.

Recall that $\hat{R}$ is an upper triangular matrix and $\tilde{Q}$ is an orthonormal matrix, (eg. $\hat{Q}^T Q = I$)

Now since A is symmetric pos. definitive, we know:

$$A = \hat{Q}\hat{R} = \hat{R}^T \hat{Q}^T = A^T$$

Now observe that

$$\hat{R} = \tilde{Q}^T A, \quad \hat{R}^T = A^T \tilde{Q}$$

$$R^T = (\hat{Q}^T A)^T = A^T (Q^T)^T = A \cdot Q$$

multiplying these gives us:

$$A^T \hat{Q} \hat{Q}^T A = \hat{R}^T \hat{R}$$

$$A^T A = \hat{R}^T \hat{R}$$, which is the cholesky factorization of $A^T A = AA$

4. (5 points) Explain how the decomposition $\mathbf{PAQ} = \mathbf{LU}$ arising in Gaussian elimination with complete pivoting (GECP) can be used to solve a linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is invertible.

observe that $PAQ = LU \implies A = P^{-1} L U Q^{-1}$

which gives us $P^{-1} L U Q^{-1} x = b$.

now let $y' = LUQ^{-1}x$, $y'' = UQ^{-1}x$ and $y''' = Qx$

we proceed as follows:

- $P^{-1} y' = b \rightarrow y' = Pb$ which is solved by multiplication
- $L y'' = y'$ which is solved by forward substitution
- $U y''' = y''$ " " " backward "
- $Q^{-1} x = y''' \rightarrow x = Qy'''$, which is solved by multiplication

This gives us the solution x to the expression $Ax = b$;
[we know such a solution exists as A is invertable.]

5. (5 points) Let $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{n \times n}$ where is invertible. Suppose that $\|\mathbf{Ay} - \mathbf{b}\|_2$ is "small". Does this imply that $\|\mathbf{y} - \mathbf{x}\|_2$ is "small" as well? Why or why not? Explain.

hm

Observe that:

$A(y-x) = Ay - Ax = Ay - b$ thus, $A^{-1}A(x-y) = A^{-1}(Ay-b) = y-x$

recall the norm inequality $\|AB\|_2 \leq \|A\|_2 \|B\|_2$; applying this gives us:

$$\|y-x\|_2 \leq \|A^{-1}\|_2 \|Ay-b\|_2$$

subsequently $\|y-x\|_2$ is only "small" if $\|A^{-1}\|_2$ is small as well.

6. (8 points) Suppose that we have computed the economic QR decomposition $\mathbf{A}_n = \hat{\mathbf{Q}}_n \hat{\mathbf{R}}_n$ where $\mathbf{A}_n = [\mathbf{a}_1, \ldots, \mathbf{a}_n] \in \mathbb{R}^{m \times n}$ ($m > n$) has full column rank. Assuming that $\mathbf{a}_{n+1} \notin \mathrm{col}(\mathbf{A}_n)$, explain how the economic QR decomposition $\mathbf{A}_{n+1} = [\mathbf{A}_n, \mathbf{a}_{n+1}] = \hat{\mathbf{Q}}_{n+1} \hat{\mathbf{R}}_{n+1}$ can be computed cheaply by making use of the previously computed factors $\hat{\mathbf{Q}}_n$ and $\hat{\mathbf{R}}_n$. *Bonus: how can this procedure be modified to still apply even when* $\mathbf{a}_{n+1} \in \mathrm{col}(\mathbf{A}_n)$*?*

CASE 1: if $a_{n+1} \notin col(A_n)$:

Recall that $\hat{Q}_{n+1}$ is a orthogonal matrix, namely all its colums be orthogonal with each other. as a_{n+1} is NOT in the colum space, it can be decomposed as:

$$a_{n+1} = \tilde{Q}_n b + v \rightarrow \tilde{Q}_n^T a_{n+1} = b + \underbrace{Q_n v}_{\rightarrow 0 \text{ as it is orthogonal}}$$

where v is the orthogonal component and $\tilde{Q}_n b$ is the component component that lies on $\tilde{Q}_n$.

Subsequently we get:

$\tilde{Q}_{n+1} = [\tilde{Q}_n, u]$, where u is the unit vector in the direction of v.

we know that $\tilde{R}_{n+1} = \begin{bmatrix} R_n & r_n \\ 0 & r_{n+1} \end{bmatrix}$, as it is an upper triangular matrix, r_n, r_{n+1} are the relevant coefficients.

Note that we need:

$$\tilde{A}_{n+1} = \underbrace{[\tilde{Q}_n, u]}_{\tilde{Q}_{n+1}} \underbrace{\begin{bmatrix} \tilde{R}_n & r_n \\ 0 & r_{n+1} \end{bmatrix}}_{\tilde{R}_{n+1}}$$

$$= \begin{bmatrix} \tilde{A}_n & \underbrace{Q_n r_n + u r_{n+1}}_{a_{n+1}} \end{bmatrix}$$

observe that $r_n = b$, found by $\tilde{Q}_n^T a_{n+1}$,

and $u r_{n+1} = v = \underbrace{a_{n+1} - \tilde{Q} b}_{\text{subtraction, multiplication.}}$

Subsequently $u = \frac{v}{\|v\|}$

and $r_{n+1} = \|v\|$

(BONUS)

case 2: when $a_{n+1} \in col(A_n)$ there is no independent vector u that would be added. Subsequently $a_{n+1} = Q_n b$.

As Q does not change, R_{n+1} simply becomes $[R_n \ b]$ to account for the new column.

Observe that

$$[Q_n][R_n \ b] = [A_n \ Q_n b]$$

$$= [A_n \ a_{n+1}] = A_{n+1}$$

7. (12 points) Part (a): Let $\mathbf{C} \in \mathbb{R}^{n\times n}$ be of the form $\mathbf{C} = \mathbf{I} + \mathbf{u}\mathbf{u}^T$ for some $\mathbf{u} \in \mathbb{R}^n$. Show that $\mathbf{C}$ is symmetric positive definite, and find an expression for $\mathbf{C}^{-1}$. *Hint:* $\mathbf{C}^{-1}$ *has the form* $\mathbf{C}^{-1} = \mathbf{I} + a\mathbf{u}\mathbf{u}^T$ *for some* $a \in \mathbb{R}$.

Part (b): Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be of the form $\mathbf{A} = \mathbf{B} + \mathbf{u}\mathbf{u}^T$ where $\mathbf{B}$ is symmetric positive definite and $\mathbf{u} \in \mathbb{R}^n$. Show that $\mathbf{A}$ can be factorized as

→ A is also symmetric pos. def. ?

→ LL^T possible..

$$\mathbf{A} = \mathbf{L}\left(\mathbf{I} + \mathbf{v}\mathbf{v}^T\right)\mathbf{L}^T \tag{5}$$

where $\mathbf{L}$ is invertible lower-triangular and $\mathbf{v} \in \mathbb{R}^n$ is some vector (to be found). u

$= LIL^T + LUU^TL^T$

$= LL^T +$

Part (c): Assume that the Cholesky factorization of $\mathbf{B}$ has already been computed. Outline an efficient procedure to solve $\mathbf{A}\mathbf{x} = \mathbf{b}$ without computing the Cholesky factorization of $\mathbf{A}$, where $\mathbf{A}$ is of the form in Part (b). How many flops does your procedure require? "Big-O" notation is fine.

PART a) Observe that for any nonzero vector $x \in \mathbb{R}^n$:

$x^T C x = x^T(I + uu^T)x$

$= \underbrace{x^T I x}_{x^Tx} + x^T u u^T x$

now observe that $x^T u = u^T x = c \in \mathbb{R}$, and $x^T x = \|x\|^2$. Plugging these in, we get:

$= \|x\|^2 + c^2 > 0$

- now observe that

$C^T = (I + uu^T)^T = I^T + (uu^T)^T$

$= I + uu^T$, as uu^T is symmetric.

$= C$

as $x^T C x > 0$ for any nonzero vector $x \in \mathbb{R}^n$, and $C^T = C$

by definition, it is symmetric positive definitive.

Let $C^{-1} = (I + auu^T)$ for some $a \in \mathbb{R}$

- Now consider:

$I = (I + auu^T)(I + uu^T)$

$I = I^2 + uu^T + auu^T + auu^Tuu^T$

- we know that $u^Tu = c$ for some $c \in \mathbb{R}$. Plugging this in:

$= I + (a+1)uu^T + au(c)u^T$

re-organizing, we get:

$= I + (a+1)uu^T + ac\,uu^T$

$= I + (a+1+ac)uu^T$

- we need $a+1+ac = 0$. The a that satisfies this is given by:

$a(1+c) = -1$

$a = \dfrac{-1}{(1+c)}$

PART b

observe that $B = LL^T$ for some lower triangular matrix L. Now observe that:

$$A = LIL^T + uu^T$$

Now consider, as L is invertible,

$$Lv = u \Rightarrow v = L^{-1}u, \quad v^T = (L^{-1}u)^T$$

This gives us: $uu^T = (Lv)(Lv)^T = Lv \cdot v^T \cdot L^T$

Plugging it back:

$$A = LIL^T + Lvv^TL^T$$

$$= L(I + vv^T)L^T$$

This proves our statement.

PART c

Recall from part b) if we know the Cholesky factorization of $B = LL^T$, we also have

$$A = L(I + vv^T)L^T \quad \text{where} \quad v = L^{-1}u.$$

Now consider $Ax = b$, let $(I + vv^T)L^T x = y'$ and $L^T x = y''$.

- $Ax = Ly' = b$ can be solved via forward subs. requiring $O(n^2)$ flops. $\Big\} O(n^2)$

- $(I + vv^T)y'' = y'$ requires computation of $Lv = u$ through forward subs. $\to O(n^2)$, addition $\to O(n)$, matrix vector multiplication $\to O(n^2)$, transposing $u \to O(1)$ $\Big\} O(n^2)$

- $L^T x = y''$ can be solved with back substitution, requiring $\to O(n^2)$ flops. $\Big\} O(n^2)$

Adding all the steps gives us $O(n^2) + O(n^2) + O(n^2) = O(n^2)$ flops for solving $Ax = b$.

8. (6 points) You have been hired as a computational scientist at `Pseudoinverse Inc.`, and your first task is to design a specialized method for computing matrix-vector products of the form $\mathbf{L}^\dagger \mathbf{x}$ given an input vector $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{L}$ is given by

$$\mathbf{L} = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}. \tag{6}$$

similar to what we had in the first labs...

Describe every numerical method that you know (number them individually, clearly) which could be used to compute these matrix-vector products. Then, choose the method you believe is the best for the job and write a short pitch ~~(2-3 sentences)~~ for your boss to read detailing why this is the best. *Hint: amongst other methods, you might consider the approximation* $\mathbf{L}^\dagger \approx (\mathbf{L}^T\mathbf{L} + \delta \mathbf{I})^{-1}\mathbf{L}^T$. *Also, depending on the argument you make, there may not be a single best method.*

1. Use $\underline{L^+ \approx (L^T L + \delta I)^{-1} L^T}$

2. Compute the $\underline{\text{QR decomposition}}$, $L = QR \rightarrow L^+ = R^{-1} Q^T$
   ↳ costly

3. $\underline{\text{SVD}}$: $L^+ = V \Sigma^+ U^T \rightarrow$ costly

4. An iterative method like gradient descent

PITCH:

The method I propose is 1! As L is a sparse matrix, the computations $L^T L$, $(L^T L + \delta I) L^T$ are not very expensive. This matters especially while we're dealing with large matrices. Furthermore as we know for $\lim_{\delta \to 0} (L^T L + \delta I)^{-1} L^T = I$, we avoid unexpected large errors!