Numerical Optimization Homework 1

James Allen

October 2, 2015

1 Problem 1

Let $P \in \mathbb{R}^{nxn}$ be a non zeros projector.

Show that $\|P\|_2 \geq 1$, and that this hold with equality if and only if P is an orthogonal projector.

From Trefethen & Bau “A Projector is a square matrix P that satisfies $P^2 = P$”

Using the Cauchy-Schwarz inequality: $|(x \cdot y)| \leq \|x\| \cdot \|y\|$ where $(j \cdot i)$ is the dot product of j and i

Now applying this to P:

$$\|(P \cdot P)\| \leq \|P\| \cdot \|P\|$$

$$\left\|P^2\right\| \leq \|P\| \cdot \|P\|$$

$$\frac{\left\|P^2\right\|}{\|P\|} \leq \|P\|$$

Since $P = P^2$

$$\cancelto{1}{\frac{\|P\|}{\|P\|}} \leq \|P\|$$

$$1 \leq \|P\|$$

*”Orthogonal projector is any projector that is Hermitian, satisfying $P^T = P$”

If the projector is symmetric and diagonalizable then you can do Eigenvalue Decomposition

$$P = Q\Lambda Q^{-1}$$

$$\left\| \left(Q\Lambda \cancelto{I}{Q^{-1} Q} \Lambda Q^{-1} \right) \right\| \leq \|Q\Lambda Q^{-1}\| \cdot \|Q\Lambda Q^{-1}\|$$

$$\left\| \left(Q \cancelto{\Lambda^2}{\Lambda\Lambda} Q^{-1} \right) \right\| \leq \|Q\Lambda Q^{-1}\| \cdot \|Q\Lambda Q^{-1}\|$$

$$\|(Q\Lambda^2 Q^{-1})\| \leq \|Q\Lambda Q^{-1}\| \cdot \|Q\Lambda Q^{-1}\|$$

This can only be true if the eigenvalues of P are 1? so $\|P\| = 1$?

2 Problem 2

Suppose that A is $m-\text{by}-n$, with $m \geq n$ and A full rank.

2.a Show that a minimizer x and corresponding residual r of the problem

$$\min_x \left\{ \frac{1}{2} \|Ax - b\|^2 \right\}$$

(where $\|\cdot\|$ is the 2-norm)

can be obtained by solving the augmented linear system

$$\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} r + Ax \\ A^T r \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

2.b Now write down the the augmented linear system that gives the solution and residual of

$$\min_x \left\{ \frac{1}{2} \|Ax - b\|^2 + \frac{1}{2} \delta^2 \|x\|^2 + c^T x \right\}$$

Expanding this

$$\frac{1}{2} x^T A^T A x - x^T A^T b + \frac{1}{2} b^T b + \frac{1}{2} \delta^2 x^T x + c^T x$$

Grouping like terms

$$\frac{1}{2} x^T \left(A^T A + \delta^2 I\right) x + \left(\left(-b^T A + c\right)^T x \right)^T + \frac{1}{2} b^T b$$

Completing the square

$$\frac{1}{2}\left\|\left(A^TA+\delta^2I\right)^{1/2}x+\left(A^TA+\delta^2I\right)^{-1}\left(A^Tb+c\right)\right\|^2$$

Showing that this actually is completing the square:

$$\frac{1}{2}\left\|\left(A^TA+\delta^2I\right)^{1/2}x+\left(A^TA+\delta^2I\right)^{-1}\left(A^Tb+c\right)\right\|^2$$

$$\frac{1}{2}x^T\left(A^TA+\delta^2I\right)x+\left(A^TA+\delta^2I\right)^{-1/2}\left(A^Tb+c\right)x+\left(A^TA+\delta^2I\right)^{-2}\left(A^Tb+c\right)^2$$

3 Problem 3

Set A by $m-by-n$, with the SVD $A = U\Sigma V^T$. Compute the SVD of the following matrices in terms of the factors of A:

Some useful notes:

$(AB)^{-1} = B^{-1}A^{-1}$

$(AB)^T = B^T A^T$

3.a $\left(A^T A\right)^{-1}$

$$\left(\left(U\Sigma V^T\right)^T U\Sigma V^T\right)^{-1}$$

Since the Left (U) and Right (V) singular vectors are orthogonal to their respective transposes:

$$\left(V\Sigma^T \cancelto{I}{U^T U}\Sigma V^T\right)^{-1}$$

$$\left(V\cancelto{\Sigma^2}{\Sigma^T \Sigma}V^T\right)^{-1}$$

As well it is known that Σ is a square diagonal matrix

$$\left(V\Sigma^2 V^T\right)^{-1}$$

$$\left(V^T\right)^{-1}\Sigma^{-2}V^{-1}$$

3.b $\left(A^T A\right)^{-1} A^T$

Carrying the simplified form of $\left(A^T A\right)^{-1}$ from the previous problem:

$$\left(V\Sigma^2 V^T\right)^{-1}\left(U\Sigma V^T\right)^T$$

simplifying the transpose:

$$\left(V\Sigma^2V^T\right)^{-1}V\Sigma^TU^T$$

factoring out the inverse

$$\left(V^T\right)^{-1}\left(\Sigma^2\right)^{-1}\cancelto{I}{V^{-1}V}\Sigma^TU^T$$

$$\left(V^T\right)^{-1}\cancelto{\Sigma^{-1}}{\left(\Sigma^2\right)^{-1}\Sigma^T}U^T$$

$$\left(V^T\right)^{-1}\Sigma^{-1}U^T$$

3.c $A\left(A^TA\right)^{-1}$

Carrying the simplified form of $\left(A^TA\right)^{-1}$ from the first problem

$$U\Sigma V^T\left(V\Sigma^2V^T\right)^{-1}$$

factoring out the inverse

$$U\Sigma\cancelto{I}{V^T\left(V^T\right)^{-1}}\left(\Sigma^2\right)^{-1}V^{-1}$$

$$U\cancelto{\Sigma^{-1}}{\Sigma\left(\Sigma^2\right)^{-1}}V^{-1}$$

$$U\Sigma^{-1}V^{-1}$$

3.d $A\left(A^TA\right)^{-1}A^T$

Carrying the simplified form of $\left(A^TA\right)^{-1}$ from the first problem

$$U\Sigma V^T\left(V\Sigma^2V^T\right)^{-1}\left(U\Sigma V^T\right)^T$$

simplifying the transpose:

$$U\Sigma V^T\left(V\Sigma^2V^T\right)^{-1}V\Sigma^TU^T$$

factoring out the inverse

$$U\Sigma \cancelto{I}{V^T \left(V^T\right)^{-1}} \left(\Sigma^2\right)^{-1} \cancelto{I}{V^{-1} V} \Sigma^T U^T$$

$$U \cancelto{I}{\Sigma \left(\Sigma^2\right)^{-1} \Sigma^T} U^T$$

$$UU^T$$

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

4 Problem 4

Suppose that the $m-\text{by}-n$ matrix, with $m<n$, is full rank. Then the problem $\min\|Ax-b\|$ is under determined.

4.a Show that the solution is an $(n-m)-$dimensional set.

$$A=\underbrace{U}_{mxm}\underbrace{\Sigma}_{mxm}\underbrace{V^T}_{mxn}$$

$$A^{\dagger}=V\Sigma U^T$$

and the solution to this is

$$A^{\dagger}\left(AA^{\dagger}\right)^{-1}$$

So we must show that

$$A^{\dagger}\left(AA^{\dagger}\right)^{-1}=A^{\dagger}$$

$$V\Sigma U^T\left(U\Sigma\cancel{V^TV}\Sigma U^T\right)^{-1}$$

$$V\Sigma U^T\left(U\Sigma^2U^T\right)^{-1}$$

$$V\Sigma\cancel{U^T\left(U^T\right)^{-1}}\Sigma^{-2}U^{-1}$$

$$\underbrace{V}_{nxm}\underbrace{\Sigma^{-1}}_{mxm}\underbrace{U^{-1}}_{mxm}=nxm$$

4.b Show how to compute the unique minimum 2-norm solution using an appropriately modified:

4.b.1 normal equations

In Dremmel:

they state that $x=\left(A^TA\right)^{-1}A^Tb$ is the minimizer of $\|Ax-b\|_2^2$, this is shown by completing the square in the text:

"$(Ax' - b)^T (Ax' - b) = (Ay + Ax - b)^T (Ay + Ax - b)$

$$(Ay)^T (Ay) + (Ax - b)^T (Ax - b) + 2(Ay)^T (Ax - b)$$

$$\|Ay\|_2^2 + \|Ax - b\|_2^2\text{''}$$

and they go on to say that this is minimized when y=0 . . . and something about orthogonal space.

4.b.2 QR decomposition

$$\left(A^T A\right)^{-1} A^T$$

and $A = QR$

$$\left(\left(R^T Q^T\right) QR\right)^{-1} \left(R^T Q^T\right)$$

$$R^{-1} Q^{-1} \cancelto{I}{Q^{-T} R^{-T} R^T Q^T}$$

$$R^{-1} Q^{-1} = (QR)^{-1}$$

4.b.3 SVD

$$\left(A^T A\right)^{-1} A^T$$

and $A = U\Sigma V^T$

$$\left(V\Sigma^2 V^T\right)^{-1} \left(U\Sigma V^T\right)^T$$

$$\left(V\Sigma^2 V^T\right)^{-1} V\Sigma^T U^T$$

$$\left(V^T\right)^{-1} \left(\Sigma^2\right)^{-1} \cancelto{I}{V^{-1} V} \Sigma^T U^T$$

$$\left(V^T\right)^{-1} \cancelto{\Sigma^{-1}}{\left(\Sigma^2\right)^{-1} \Sigma^T} U^T$$

$$\left(V^T\right)^{-1} \Sigma^{-1} U^T$$

5 Problem 5

https://github.com/jamestallen/MAT258A/blob/master/HW_1_Plots.ipynb