MATH 270A – NUMERICAL LINEAR ALGEBRA – HOMEWORK 3

Due Friday, November 9th. Handwritten submissions only.

Exercise 1
Given a vector $v \in \mathbb{R}^2$, we want to define a rotation matrix $G \in \mathbb{R}^{2\times 2}$ and $\gamma \in \mathbb{R}$ such that $Gv = \gamma e^1$. Write

$$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}.$$

Show that the following two constructions provide the coefficients c and s and the value γ.

- Variant 1:

$$\gamma = v_2 \cdot \nu, \quad s = \frac{1}{\nu}, \quad c := s\tau, \quad \tau = \frac{v_1}{v_2}, \quad \nu = \sqrt{1+\tau^2}.$$

- Variant 2:

$$\gamma = v_1 \cdot \nu, \quad c = \frac{1}{\nu}, \quad s := c\tau, \quad \tau = \frac{v_2}{v_1}, \quad \nu = \sqrt{1+\tau^2}.$$

Solution 1
Depending on the whether first or second variant is used, we have either of

$$\nu = \sqrt{1 + \left(\frac{v_1}{v_2}\right)^2} = \frac{1}{v_2}\sqrt{v_2^2 + v_1^2} = \frac{1}{v_2}\|v\|_2,$$
$$\nu = \sqrt{1 + \left(\frac{v_2}{v_1}\right)^2} = \frac{1}{v_1}\sqrt{v_1^2 + v_2^2} = \frac{1}{v_1}\|v\|_2.$$

In either case, it follows that $\gamma = \|v\|_2$. Now, in the first variant we have

$$s = \frac{1}{\nu} = \frac{v_2}{\|v\|_2} = \sin(\alpha),$$
$$c = s\tau = \frac{1}{\nu}\frac{v_1}{v_2} = \frac{v_2}{\|v\|_2}\frac{v_1}{v_2} = \frac{v_1}{\|v\|_2} = \cos(\alpha),$$

where α is the angle between v and first unit vector. In the second variant, we have

$$c = \frac{1}{\nu} = \frac{v_1}{\|v\|_2} = \cos(\alpha),$$
$$s = c\tau = \frac{1}{\nu}\frac{v_2}{v_1} = \frac{v_1}{\|v\|_2}\frac{v_2}{v_1} = \frac{v_2}{\|v\|_2} = \sin(\alpha).$$

By the definition of the Givens rotation, the matrix rotates every vector by the angle $-\alpha$. In particular, v is rotated onto a multiple of the first unit vector. This completes the proof.

Remark 1
In applications, the first variant is prefered if $|v_2| \geq |v_1|$ and the second variant is prefered if $|v_1| \geq |v_2|$.

Exercise 2
Let $A \in \mathbb{R}^{n\times m}$ with $\ker(A) = 0$ and let $b \in \mathbb{R}^n$. Show that the matrix

$$M = \begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} \tag{1}$$

is invertible. Prove that the solution of the system

$$\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} r \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \tag{2}$$

satisfies $A^T Ax = A^T b$.

Solution 2
Assume that $(r, x) \in \mathbb{R}^{n+m}$ satisfies $A^T r = 0$ and $r + Ax = 0$. Then $r \perp \operatorname{ran}(A)$. Since $Ax \in \operatorname{ran}(A)$, we have $r = 0$ and $Ax = 0$. Hence $x = 0$. Since M is square, it must be invertible.

The system in M has a unique solution. Let x be the unique solution of $A^T Ax = A^T b$. Write $r = b - Ax$. Then (r, x) is the solution of the system in M.

Exercise 3
Let $A \in \mathbb{R}^{n\times m}$ with $\operatorname{ran}(A) = \mathbb{R}^n$ and let $b \in \mathbb{R}^n$. Show that the matrix

$$M = \begin{pmatrix} -I & A^T \\ A & 0 \end{pmatrix} \tag{3}$$

is invertible. Prove that the solution of the system

$$\begin{pmatrix} -I & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} \tag{4}$$

satisfies $Ax = b$ with $x = A^T y$.

Solution 3
Let $x \in \mathbb{R}^{(m)}$ with $Ax = b$ and $\ker(A) \perp x$. Then there exists $y \in \mathbb{R}^n$ with $x = A^T y$. The vector y is unique. So the system in M has a solution, and since M is square, M is invertible.

The rest is trivial.

Exercise 4
Let $D \in \mathbb{R}^{n\times m}$ and $T \in \mathbb{R}^{m\times n}$ be matrices whose entries are zero except for the first d diagonal entries $D_{11}, D_{22}, \ldots, D_{dd}$ and $T_{11}, T_{22}, \ldots, T_{dd}$, respectively, where $d \le \min(n, m)$. Assume that $T_{ii} = D_{ii}^{-1}$ for $1 \le i \le d$. Show that $T = D^\dagger$.

Solution 4
Let D and T as in the statement of the exercise. We show that for any $y \in \mathbb{R}^n$ we have $x := Ty \in \mathbb{R}^{(m)}$ being the solution of the generalized least-squares problem.

We have $\ker(D)^\perp$ and $\operatorname{ran}(D)$ being spanned by the first d standard unit vectors in $\mathbb{R}^{(m)}$ and $\mathbb{R}^n$, respectively.

Let $y \in \mathbb{R}^n$. We have $d \le m$ and $d \le n$. Let $x = Ty$. A direct computation shows that $x \in \ker(D)^\perp$ and that $(Dx)_i = y_i$ for $1 \le i \le d$, so $y - Dx \in \operatorname{ran}(A)^\perp$. This shows that x is the least squares solution, and we conclude that T must be the pseudoinverse of D.

Exercise 5
Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{m\times n}$. Show that $B = A^\dagger$ if and only if the following four conditons hold:

$$ABA = A, \quad BAB = B, \quad (AB)^T = AB, \quad (BA)^T = BA.$$

Proof. Let $A = WDV^T$ be the (full) singular value decomposition of A. From the lecture and the previous exercise, we know that

$$A^\dagger = VD^\dagger W.$$

We verify that the pseudoinverse $A^\dagger$ satisfies the four required identities by simply plugging in the pseudoinverse and doing elementary computations.

We show that the pseudoinverse $A^\dagger$ is the only solution to the four equations

$$\begin{aligned} ABA &= A, \\ BAB &= B, \\ (AB)^T &= AB, \\ (BA)^T &= BA. \end{aligned}$$

One easily checks that $(A^T)^\dagger = (A^\dagger)^T$ using the singular value decomposition. Now we find

$$A^\dagger AB = A^\dagger ABAB = A^\dagger AA^T B^T B = A^T (A^\dagger)^T A^T B^T B = A^T B^T B = (BA)^T B = BAB = B$$

Taking the transpose of the first two conditions, we get $A^T B^T A^T = A^T$ and $B^T A^T B^T = B^T$. Now we find

$$A^\dagger AB = A^\dagger B^T A^T = A^\dagger AA^\dagger B^T A^T = A^\dagger A^{T\dagger} A^T B^T A^T = A^\dagger A^{T\dagger} A^T = A^\dagger AA^\dagger = A^\dagger.$$

This completes the proof. □

Exercise 6
Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction with contraction constant $q \in [0, 1)$. Let $x^* \in \mathbb{R}^n$ be the unique fixpoint of T. Let $x^{(0)} \in \mathbb{R}^n$ and define recursive $x^{(m+1)} = T(x^{(m)})$ for $m \in \mathbb{N}_0$.

Let $\epsilon > 0$. Find the minimum number M for which $\|x^{(M)} - x^*\| < \epsilon$ in terms of ϵ and

(a) the initial error norm $\|x^0 - x^*\|$.
(b) the initial difference $\|x^0 - x^1\|$.

Solution 5
The calculations are very similar in each case.

- A simple induction argument shows that

  $$\|x^{(M)} - x^*\| \le q^M \|x^{(0)} - x^*\|.$$

  Assuming $q \ne 0$, we have

  $$\begin{aligned} &q^M \|x^{(0)} - x^*\| < \epsilon \\ &\iff q^M < \frac{\epsilon}{\|x^{(0)} - x^*\|} \\ &\iff M > \log_q \left(\frac{\epsilon}{\|x^{(0)} - x^*\|} \right). \end{aligned}$$

- From the lecture, we know that

  $$\|x^{(M)} - x^*\| \le \frac{q^M}{1-q} \|x^{(0)} - x^*\|.$$

Similar as before, we find

$$\begin{aligned}
&\frac{q^M}{1-q}\|x^{(1)}-x^{(0)}\| < \epsilon \\
\iff &\frac{q^M}{1-q} < \frac{\epsilon}{\|x^{(1)}-x^{(0)}\|} \\
\iff &q^M < (1-q)\frac{\epsilon}{\|x^{(1)}-x^{(0)}\|} \\
\iff &M > \log_q\left((1-q)\frac{\epsilon}{\|x^{(1)}-x^{(0)}\|}\right).
\end{aligned}$$

The proof is complete.

Exercise 7
Consider the fixpoint iteration

$$T(x) = x + P^{-1}(b - Ax).$$

Let $x^{(0)} \in \mathbb{R}^n$ and define recursive $x^{(m+1)} = T(x^{(m)})$ for $m \in \mathbb{N}_0$. For each $m \in \mathbb{N}_0$ we define the m-th residual as

$$r^{(m)} = b - Ax^{(m)}.$$

Show that for all $m \in \mathbb{N}_0$ we have

$$r^{(m+1)} := (I - AP^{-1})r^{(m)}.$$

Solution 6
Let $m \in \mathbb{N}_0$. Then we have

$$\begin{aligned}
r^{(m+1)} &= b - Ax^{(m+1)} \\
&= b - A\left(x^{(m)} + P^{-1}\left(b - Ax^{(m)}\right)\right) \\
&= b - Ax^{(m)} + AP^{-1}\left(b - Ax^{(m)}\right) \\
&= r^{(m)} + AP^{-1}r^{(m)}.
\end{aligned}$$

This had to be shown.

Exercise 8
Consider the matrix $A \in \mathbb{R}^{2\times 2}$ and vectors $x, b \in \mathbb{R}^2$ with $Ax = b$, where

$$A = \begin{pmatrix} 11 & -1 \\ -1 & 9 \end{pmatrix}, \quad x = \begin{pmatrix} 5 \\ -2 \end{pmatrix}, \quad b = \begin{pmatrix} 57 \\ -23 \end{pmatrix}. \tag{5}$$

We guess that the diagonal matrix $P^{-1} \in \mathbb{R}^{2\times 2}$ with

$$P^{-1} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix} \tag{6}$$

is a good approximate inverse of A. Conduct the iteration

$$x^{(k+1)} = x^{(k)} + P^{-1}(b - Ax^{(k)}) \tag{7}$$

with starting value $x^{(0)} = 0$ for six steps and calculate $\|x - x^k\|$ at each step. Compute the eigenvalues of $I - P^{-1}A$ and compare with the ratio of the norms of consecutive errors $\|x - x^{k+1}\|/\|x - x^k\|$.

You can use Matlab or GNU Octave to solve this problem. If you don't have access to Matlab or GNU Octave, you can use `https://octave-online.net/` for this exercise.

Solution 7

A simple computation shows

step	approximate solution	error	error norm	ratio
0	$(0,0)^T$	$(5,-2)^T$	$\sqrt{29}$	–
1	$(5.7,-2.3)^T$	$(0.7,0.3)^T$	0.76158	0.14142
2	$(4.9,-1.96)^T$	$(0.1,-0.04)^T$	0.1077	0.14142
3	$(5.014,-2.006)^T$	$(-0.014,0.06)^T$	0.015232	0.14142
4	$(4.9980,-1.9992)^T$	$\left(2\times10^{-3},-8\times10^{-4}\right)^T$	0.0021541	0.14142
5	$(5.0003,-2.0001)^T$	$\left(-2.8\times10^{-4},1.2\times10^{-4}\right)^T$	3.0463×10^{-4}	0.14142
6	$(5,-2)^T$	$\left(4\times10^{-5},-1.6\times10^{-5}\right)^T$	4.3081×10^{-5}	0.14142

One finds that ±0.14142 are the eigenvalues of the iteration matrix.