Math 501 **Name:** Qilin Ye
Spring 2021
Midterm (take-home)
00:01 AM Fri, Mar. 19 – 11:59 PM Sun, Mar. 21.
Time: Submit by 11:59 PM on Sunday, Mar. 21

This exam contains 14 pages (including this cover page) and 9 questions. Total of points is 165.

- **Show your work.**
- **Finish by yourself!**

Grade Table (for teacher use only)

Question	Points	Score
1	10	
2	20	
3	30	
4	25	
5	15	
6	20	
7	15	
8	10	
9	20	
Total:	165	

MATH 501 Midterm

Qilin Ye

March 20, 2021

Problem 1

The exponential integrals are the functions E_n defined by

$$E_n(x) = \int_1^\infty (e^{xt}t^n)^{-1}\,\mathrm{d}t, \quad (n \geqslant 0, x > 0).$$

These functions satisfy the equation

$$nE_{n+1}(x) = e^{-x} - xE_n(x).$$

If $E_1(x)$ is known, can this computation be used to compute $E_2(x), E_3(x), \ldots$ accurately?

Solution

In general, no. We first rewrite the equation as

$$E_{n+1}(x) = \frac{e^{-x}}{n} - \frac{x}{n}E_n(x).$$

This is relatively accurate for small x. When x is small, $x/n < 1$ for almost n's and the error of δ units in $E_n(x)$ gradually decays to 0.

For larger x's, however, every time when we compute $E_{n+1}(x)$ for $n < x$, the error is multiplied by a factor > 1 as $x/n > 1$ when $n < x$. For example, when $x = 10$, the initial error δ caused by rounding $E_1(x)$ becomes a stunning

$$\frac{10}{1}\cdot\frac{10}{2}\cdot\cdots\cdot\frac{10}{9}\cdot\delta = \frac{10^9\delta}{9!} \approx 2755\delta,$$

which greatly disrupts the accuracy. In fact, when trying $x = 15$ on my MATLAB with a machine epsilon 2^{-52}, $E_{12}(x)$ returns a negative value, which is absurd.

Problem 2 has been deleted; see problem 5 instead.

Problem 3: (a)

(Problem 4.7.1, p.217) Prove that if A is symmetric, then the gradient of the function

$$q(x) = \langle x, Ax\rangle - 2\langle x, b\rangle$$

at x is $2(Ax - b)$.

Proof. First notice that the partial derivative

$$g := \frac{\partial\langle x, y\rangle}{\partial x} = y. \tag{1}$$

Indeed, consider g_1 (the first component of g). It is defined as

$$\frac{\partial}{\partial x_1}\sum_{i=1}^{n} x_i y_i = y_1,$$

and likewise for all the remaining ones. Vector-wise we also have the chain rule:

$$\begin{aligned}\frac{\mathrm{d}\langle x, Ax\rangle}{\mathrm{d}x} &= \frac{\mathrm{d}x^T}{\mathrm{d}x}\cdot\frac{\partial\langle x, Ax\rangle}{\partial x} + \frac{\mathrm{d}(Ax)^T}{\mathrm{d}x}\cdot\frac{\partial\langle x, Ax\rangle}{\partial(Ax)}\\ &= Ax + \frac{\mathrm{d}(x^T A^T)}{\mathrm{d}x}\cdot x\\ &= Ax + A^T x = (A + A^T)x = 2Ax.\end{aligned}$$

From (1), it is also clear that $\partial(2\langle x, b\rangle)/\partial x = 2b$. Therefore the gradient of $q(x) = 2(Ax - b)$. □

Problem 3: (b)

(Problem 4.7.6, p.218) Prove that if $\hat{t}$ is defined by

$$\hat{t} = \frac{\langle v, b - Ax\rangle}{\langle v, Av\rangle}$$

and if $y = x + \hat{t}v$, then $v \perp (b - Ay)$, that is, $\langle v, b - Ay\rangle = 0$.

Proof. Indeed, this result follows directly from brute-force computation:

$$\begin{aligned}\langle v, b - Ay\rangle &= \langle v, b - Ax\rangle - \left\langle v, A\hat{t}v\right\rangle\\ &= \frac{\langle v, b - Ax\rangle}{\langle v, Av\rangle}\cdot\langle v, Av\rangle - \left\langle v, A\hat{t}v\right\rangle\\ &= \hat{t}\langle v, Av\rangle - \left\langle v, A\hat{t}v\right\rangle = \left\langle v, A\hat{t}v\right\rangle - \left\langle v, A\hat{t}v\right\rangle = 0.\end{aligned}$$

□

Problem 3: (c)

(Problem 4.7.7, p.218) Show that in the method of steepest descent,

$$q(x^{(k+1)}) = q(x^{(k)}) - \|r^{(k)}\|^4 / \left\langle r^{(k)}, Ar^{(k)}\right\rangle$$

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

Proof. Recall (as shown in the textbook) that

$$q(x + tv) = q(x) + 2t \langle v, Ax - b \rangle + t^2 \langle v, Av \rangle .$$

Here we simply need to plug in the variables. Notice that $x^{(k+1)} = x^{(k)} + t_k r^{(k)}$, so

$$\begin{aligned} q(x^{(k+1)}) &= q(x^{(k)}) + 2t_k \left\langle r^{(k)}, Ax - b \right\rangle + t_k^2 \left\langle r^{(k)}, Ar^{(k)} \right\rangle \\ &= q(x^{(k)}) - 2t_k \left\langle r^{(k)}, r^{(k)} \right\rangle + t_k^2 \left\langle r^{(k)}, Ar^{(k)} \right\rangle \\ &= q(x^{(k)}) - 2 \cdot \frac{\left\langle r^{(k)}, r^{(k)} \right\rangle}{\left\langle r^{(k)}, Ar^{(k)} \right\rangle} \cdot \left\langle r^{(k)}, r^{(k)} \right\rangle + \frac{\left\langle r^{(k)}, r^{(k)} \right\rangle^2}{\left\langle r^{(k)}, Ar^{(k)} \right\rangle^2} \left\langle r^{(k)}, Ar^{(k)} \right\rangle \\ &= q(x^{(k)}) - \frac{\left\langle r^{(k)}, r^{(k)} \right\rangle^2}{\left\langle r^{(k)}, Ar^{(k)} \right\rangle} \\ &= q(r^{(k)}) - \| r^{(k)} \|^4 / \left\langle r^{(k)}, Ar^{(k)} \right\rangle . \end{aligned}$$

□

Problem 4: (a)

Consider an iterative method $x^{(k+1)} = Bx^{(k)} + c$, for which $\|B\| \leqslant \beta < 1$. Show that if $\|x^{(k)} - x^{(k-1)}\| < \epsilon(1-\beta)/\beta$ for all k, then $\|x - x^{(k)}\| \leqslant \epsilon$, where x is the fixed point of the iterative method.

Proof. Notice that

$$\begin{aligned} \|x^{(k)} - x^{(k-1)}\| &= \|Bx^{(k-1)} + c - Bx^{(k-2)} - c\| = \|Bx^{(k-1)} - Bx^{(k-2)}\| \\ &\leqslant \|B\| \|x^{(k-1)} - x^{(k-2)}\| \leqslant \beta \|x^{(k-1)} - x^{(k-2)}\|. \end{aligned}$$

Therefore,

$$\begin{aligned} \|x - x^{(k)}\| &= \left\| \sum_{i=k+1}^{\infty} x^{(i)} - x^{(i-1)} \right\| \leqslant \sum_{i=k+1}^{\infty} \|x^{(i)} - x^{(i-1)}\| \\ &\leqslant \sum_{i=1}^{n} \beta^i \|x^{(k+i)} - x^{(k+i-1)}\| \leqslant \frac{\beta}{1-\beta} \cdot \frac{\epsilon(1-\beta)}{\beta} = \epsilon. \end{aligned}$$

□

Problem 4: (b)

Show that the infinite series

$$I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots$$

converges for any square matrix A. Denote the sum of the series by e^A. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A, repeated according to their multiplicity, and show that the eigenvalues of e^A are $e^{\lambda_1}, \ldots, e^{\lambda_n}$.

Proof. Let a subordinated $\|\cdot\|$ be given. Recall that $\|A^n\| \leqslant \|A\|^n$. Therefore,

$$\left\| \sum_{k=0}^{n} \frac{A^k}{k!} \right\| \leqslant \sum_{k=0}^{n} \|A^k / k!\| \leqslant \sum_{k=0}^{n} \|A\|^k / k! \to e^{\|A\|}.$$

This shows that indeed the infinite series converges. Let λ_1 be an eigenvalue of A. It follows that λ_1^2 is an eigenvalue of A^2 with the same eigenvector x_1:

$$A^2 x_1 = A(Ax_1) = A(\lambda_1 x_1) = \lambda(Ax_1) = \lambda_1^2 x_1.$$

Inductively one can show that λ_1^n is an eigenvalue of A^n (and with eigenvector x_1). Therefore,

$$e^A(x_1) = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!}\right)(x_1) = \left(\sum_{k=0}^{\infty} \frac{\lambda_1^k}{k!}\right)(v) = e^{\lambda_1}(v).$$

The proof that λ_i corresponds to e^{λ_i} is simply analogous and is omitted. □

Problem 5

Show that

$$x_{n+1} = \frac{x_n(x_n^2 + 3a)}{3x_n^2 + a}, \quad n \geqslant 0$$

is a third-order method for computing $\sqrt{a}$. Calculate

$$\lim_{n\to\infty} \frac{\sqrt{a} - x_{n+1}}{(\sqrt{a} - x_n)^3}$$

assuming x_0 is chosen sufficiently close to $\sqrt{a}$.

Proof. Let e_n be the error $\sqrt{a} - x_n$. It follows that

$$\begin{aligned} e_{n+1} = \sqrt{a} - x_{n+1} &= \sqrt{a} - \frac{x_n(x_n^2 + 3a)}{3x_n^2 + a} \\ &= \frac{\sqrt{a}^3 - 3\sqrt{a}^2 x_n + 3\sqrt{a} x_n^2 - 3a^3}{3x_n^2 + a} \\ &= \frac{e_n^3}{3x_n^2 + a}. \end{aligned}$$

Therefore the convergence is of third-order. Assuming x_0 is sufficiently close to $\sqrt{a}$ so that $x_n \to \sqrt{a}$, we have

$$\lim_{n\to\infty} \frac{\sqrt{a} - x_{n+1}}{(\sqrt{a} - x_n)^3} = \lim_{n\to\infty} \frac{e_{n+1}}{e_n^3} = \lim_{n\to\infty} \frac{1}{3x_n^2 + a} = \frac{1}{4a}.$$ □

Problem 6: (a)

Let $A_{n\times n}$ be a real, symmetric, positive definite matrix. Prove that

$$|\det(A)| \leqslant \prod_{i=1}^{n} a_{i,i},$$

where the equality holds only if A is diagonal.

Proof. If A is positive definite then it is invertible, as $x^T Ax > 0$ for all nonzero x implies that $Ax \neq 0$ for all nonzero x. Recall that A admits a Cholesky factorization $A = LL^2$ and $\det(A) = \det(L)\det(L^T) = \det(L^2)$. Therefore, it suffices to prove that

$$\det(L)^2 \leqslant \prod_{i=1}^{n} a_{i,i}.$$

Notice that $\det(L)^2$ is simply the square product of the diagonal entries of L, i.e.,

$$\det(L)^2 = \prod_{i=1}^{n} L_{i,i}^2.$$

On the other hand, $a_{i,i} = \|L_i\|_2^2$, where L_i denotes the i^{th} row of L. In other words,

$$a_{i,i} = \sum_{j=1}^{i} L_{i,j}^2.$$

Since squares are nonnegative, we indeed have

$$L_{i,i}^2 \leqslant a_{i,i} \implies \det(L)^2 = \det(A) \leqslant \prod_{i=1}^{n} a_{i,i}.$$

(Notice that $\det(A) > 0$ and we may drop the absolute value.) It is also clear that if $\det(A) = \prod_{i=1}^{n} a_{i,i}$ then each $a_{i,i}$ has to equal to $L_{i,i}^2$. Therefore, L must not contain nonzero off-diagonal entries, i.e., L is diagonal. Then A is diagonal, and this proves the second claim. □

Problem 6: (b)

Show that the matrix

$$A = \begin{bmatrix} a & -1 & \cdots & -1 \\ -1 & a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & a \end{bmatrix}$$

is positive definite if $a > \sqrt{n}$.

Proof. I tried to use the hint and obtain a Cholesky factorization of A. I know that, if I could find such L then $x^T Ax = x^T LL^T x = \langle L^T x, L^T x\rangle > 0$ as long as $x \neq 0$. However, I was not able to prove the existence of L. Instead, this proof consists of brute force computation. Let $x := (x_1, \dots, x_n)^T \in \mathbb{R}^n$ be given. It follows that

$$Ax = \begin{bmatrix} ax_1 - \sum_{i=2}^n x_n & -x_1 + ax_2 & -x_1 + ax_3 & \cdots & -x_1 + ax_n \end{bmatrix}^T$$

and

$$\begin{aligned} x^T Ax &= ax_1^2 - \sum_{i=2}^{n} x_1 x_i + \sum_{j=2}^{n} (ax_j^2 - x_1 x_j) \\ &= ax_1^2 - 2\sum_{i=2}^{n} x_1 x_i + \sum_{j=2}^{n} ax_j^2 \\ &= \left(a - \frac{n-1}{a}\right) x_1^2 + \sum_{i=2}^{n} \frac{x_1^2}{a} - 2\sum_{i=2}^{n} x_1 x_i + \sum_{i=2}^{n} ax_i^2 \\ &= \left(a - \frac{n-1}{a}\right) x_1^2 + \sum_{i=2}^{n} \left(\frac{1}{\sqrt{a}} x_1 + \sqrt{a} x_i\right)^2. \end{aligned}$$

From this, we see that in fact $a - (n-1)/a > 0$ is a *sufficient* condition. In other words, if $a > \sqrt{n-1}$ then A is positive definite. Our problem asks for something slightly weaker but it of course holds too. *I have implemented Cholesky factorization of matrices of this form in Matlab; indeed, given n, if $a < \sqrt{n-1}$ then no (real) Cholesky factorization exists and if $a > \sqrt{(n-1)}$ then all good.* □

Problem 7

Using Q as in the Gauß-Seidel method, prove if A is diagonally dominant then $\|I - Q^{-1}A\|_\infty < 1$.

Proof. Before going straight into the main proof, we begin by deriving two inequalities and one equation which will be enormously helpful. Recall that A is diagonally dominant, i.e.,

$$|a_{i,i}| > \sum_{j\neq i}|a_{i,j}| = \sum_{i<i}|a_{i,j}| + \sum_{j>i}|a_{i,j}|.$$

This leads to

$$|a_{i,i}| - \sum_{j<i}|a_{i,j}| > \sum_{j>i}|a_{i,j}| \text{ for all } 1 \leqslant i \leqslant n$$

and so

$$\frac{\sum_{j>i}|a_{i,j}|}{|a_{i,i}| - \sum_{j<i}|a_{i,j}|} \leqslant \alpha < 1 \text{ for all } 1 \leqslant i \leqslant n \text{ and some } \alpha \in \mathbb{R}. \tag{1}$$

Also, since Q in the Gauß-Seidel method is simply the lower triangular matrix (with diagonal), we can write A as $Q + U$ where U is purely upper triangular, i.e., diagonal entries being uniformly 0. Then,

$$I - Q^{-1}A = I - Q^{-1}(Q + U) = I - Q^{-1}Q - QU = -QU$$

and so

$$\|I - Q^{-1}A\| = \|Q^{-1}U\|. \tag{2}$$

One more triangle inequality trick before we proceed to our main proof:

$$|a + b - b| \leqslant |a + b| + |-b| \implies |a| \leqslant |a + b| + |b|$$

so

$$|a + b| \geqslant |a| - |b|. \tag{3}$$

Main proof. Define $y = y(x) \coloneqq Q^{-1}Ux$ where x is any arbitrary vector with $\|x\|_\infty = 1$. We want to show that there exists some $k < 1$ such that $\|y\|_\infty < \beta$ holds for all x. Assume $\|y\|_\infty = |y_k|$ with $1 \leqslant k \leqslant n$. Now we consider the k^{th} component of (Qy), namely $(Qy)_k$:

$$\begin{aligned}
|(Qy)_k| &= \Big|\sum_{j\leqslant k} a_{k,j}y_j\Big| \geqslant |a_{k,k}y_k| - \Big|\sum_{j<k} a_{k,j}y_j\Big| && \text{(by (3))}\\
&\geqslant |a_{k,k}|\|y\|_\infty - \sum_{j<k}|a_{k,j}||y_j| \geqslant |a_{k,k}|\|y\|_\infty - \sum_{j<k}|a_{k,j}|\|y\|_\infty\\
&= \Big(|a_{k,k}| - \sum_{j<k}|a_{k,j}|\Big)\|y\|_\infty. && (4)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
|(Qy)_k| &= |(QQ^{-1}Ux)_k| = |(Ux)_k|\\
&= \Big|\sum_{j>k} a_{k,j}x_j\Big| \leqslant \sum_{j>k}|a_{k,j}||x_j|\\
&\leqslant \Big(\sum_{j>k}|a_{k,j}|\Big)\|x\|_\infty = \Big(\sum_{j>k}|a_{k,j}|\Big). && (5)
\end{aligned}$$

Combining (4) and (5) (and using (1)), we see

$$\|y\|_\infty \leqslant \frac{\sum_{j>k}|a_{k,j}|}{|a_{k,k}| - \sum_{j<k}|a_{k,j}|} \leqslant \beta < 1,$$

so $\|Q^{-1}U\|_\infty = \sup_{\|x\|_\infty=1} \|Q^{-1}Ux\| = \sup \|y\|_\infty \leqslant \beta < 1$, and the claim follows from (2). □

Problem 8

Is there a matrix A such that $\rho(A) < \|A\|$ for all subordinate matrix norms? Prove or disprove.

Solution

Yes! Consider $A := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, of which both eigenvalues are 0 and thus $\rho(A) = 0$. However, since A is not the zero matrix, all norms (in particular, the subordinate ones) satisfy $\|A\| > 0$.

Problem 9: (a)

The linear system

$$\begin{bmatrix} 1 & -a \\ -a & 1 \end{bmatrix} x = b$$

where a is real can under certain conditions be solved by SOR

$$\begin{bmatrix} 1 & 0 \\ -\omega a & 1 \end{bmatrix} x^{(k+1)} = \begin{bmatrix} 1-\omega & \omega a \\ 0 & 1-\omega \end{bmatrix} x^{(k)} + \omega b.$$

The corresponding iteration matrix B_ω is

$$\begin{bmatrix} 1 & 0 \\ -\omega a & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1-\omega & \omega a \\ 0 & 1-\omega \end{bmatrix}.$$

Show that the eigenvalues of the corresponding iteration matrix B_ω satisfies the following:

$$\lambda^2 - \lambda\left[2(1-\omega) + \omega^2 a^2\right] + (1-\omega)^2 = 0.$$

Proof. This can be shown by brute force computation. First notice that

$$\begin{bmatrix} 1 & 0 \\ -\omega a & 1 \end{bmatrix}^{-1} = \frac{1}{1+\omega a} \begin{bmatrix} 1 & 0 \\ \omega a & 1 \end{bmatrix}.$$

The fraction $1/(1+\omega a)$ is just a constant; the eigenvalues of B_ω and of $(1+\omega a)B_\omega$ are the same, so for convenience we drop this fraction. Then,

$$\begin{bmatrix} 1 & 0 \\ \omega & 1 \end{bmatrix} \begin{bmatrix} 1-\omega & \omega a \\ 0 & 1-\omega \end{bmatrix} = \begin{bmatrix} 1-\omega & \omega a \\ \omega a(1-\omega) & \omega^2 a^2 + 1 - \omega \end{bmatrix}.$$

To find the eigenvalues, we suppose $\det(B_\omega - \lambda I) = 0$, that is,

$$\begin{aligned} 0 = \det(B_\omega - \lambda I) &= \begin{vmatrix} (1-\omega)-\lambda & \omega a \\ \omega a(1-\omega) & (\omega^2 a^2 + 1 - \omega) - \lambda \end{vmatrix} \\ &= \left[(1-\omega)-\lambda\right]\left[(\omega^2a^2+1-\omega)-\lambda\right] - \omega^2 a^2(1-\omega) \\ &= (1-\omega)(\omega^2a^2+1-\omega) + \lambda^2 - \lambda\left[2(1-\omega)+\omega^2a^2\right] - \omega^2a^2(1-\omega) \\ &= \lambda^2 - \lambda\left[2(1-\omega)+\omega^2a^2\right] + (1-\omega)^2. \end{aligned}$$

□

Problem 9: (b)

Show that for $\omega = 1$ the eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = |a^2|$. Conclude that for this choice the iteration converges for $|a| < 1$.

Proof. From the characteristic polynomial derived in (a), if $\omega = 1$ then

$$\lambda^2 - a^2\lambda = 0 \implies \lambda_1 = 0, \lambda_2 = a^2|a|^2 \quad (a \in \mathbb{R} \text{ by assumption}).$$

Indeed, if $|a| < 1$ then $\rho(B_\omega) < 1$ and thus the iterative method

$$x^{(k)} = B_\omega x^{(k-1)} + c$$

converges by one of the theorems shown in class just recently. □