MATH 714 HOMEWORK 2 PROBLEM B AND C

Kangqi Fu
AMEP
University of Wisconsin Madison
Madison, WI 53706
kfu9@wisc.edu

November 2, 2020

ABSTRACT

This is the solution of problem B and problem C in Homework 2. Problem A is very long so I decided to write the solution by hand. It is attached at the end of this file. Thank you for the help from Diya Yang and Lewis Gross for discussing some of the homework problem with me. Without their help, I wouldn't have the chance to learn more than just finishing this problem set

1 Problem B

The code to run for this problem is in github page and the file name is Hw2B.m. The analysis here is very simple, basically, you need to assign the number of query point large enough and based on the equation, this value should be larger than or equal to around 550, because after that the number of N is stable to 100. Therefore, the critical value of N is 100. Below is the brief code:

```
N = 0;
dif = 9999;
xq = 0:0.0001:1;
while dif > 0.01
    N = N+1;
    x = linspace(0,1,N+1);
    v = exp(-400*(x-0.5).^2);
    vq = interp1(x,v,xq);
    f = exp(-400*(xq-0.5).^2);
    dif = norm(f-vq,Inf);
end
disp(['The smallest value of N = ', num2str(N)]);
```

2 Problem C

2.1 Question (a)

Based on the information provided, the numerical scheme is:

$$\frac{u_{i,j}^{m-1} - 2u_{i,j}^{m} + u_{i,j}^{m+1}}{\Delta t^2} = \frac{u_{i-1,j}^{m} - 2u_{i,j}^{m} + u_{i+1,j}^{m}}{\Delta x^2} + \frac{u_{i,j-1}^{m} - 2u_{i,j}^{m} + u_{i,j+1}^{m}}{\Delta y^2} \tag{1}$$

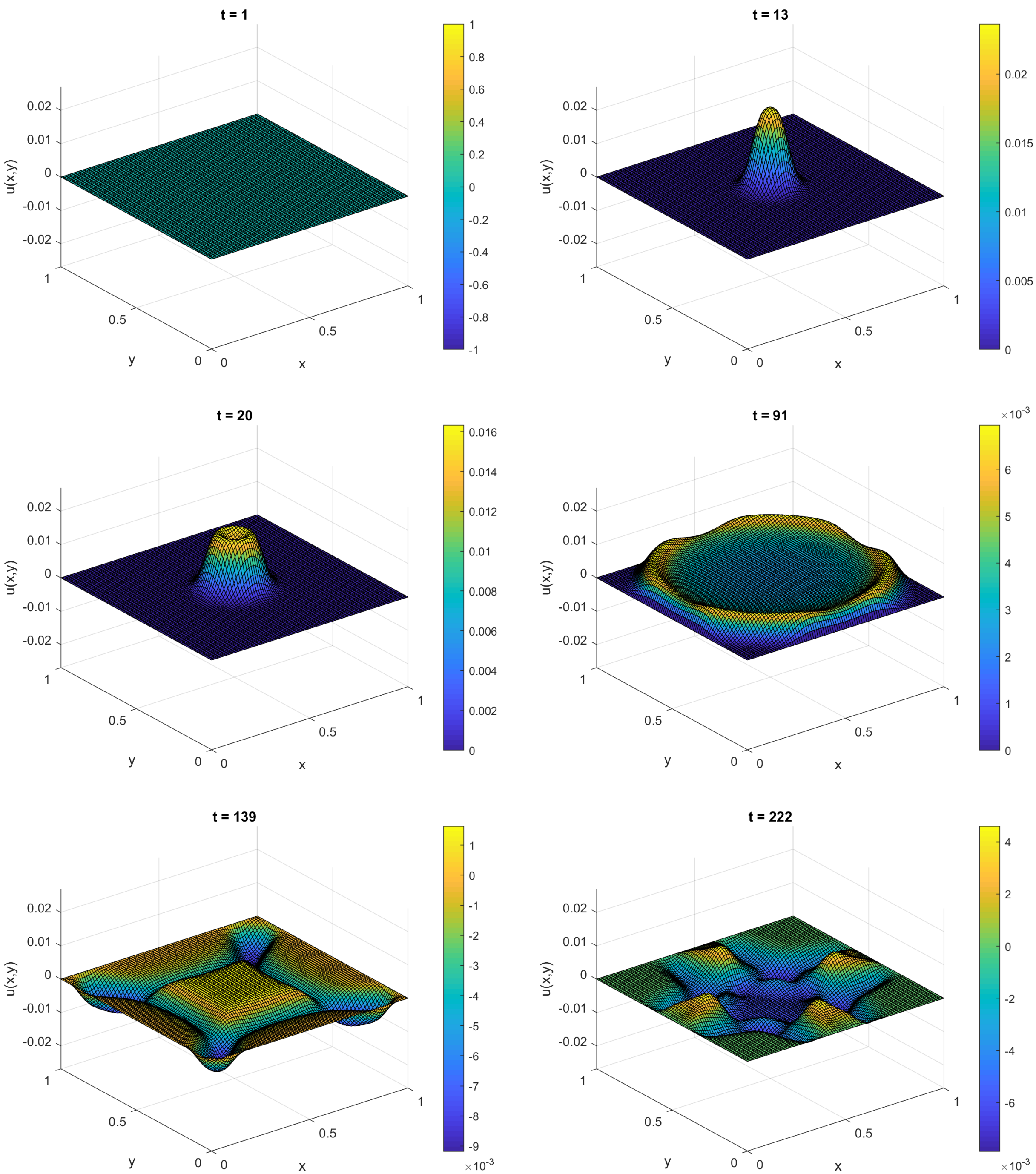
With $\Delta x^2 = \Delta y^2$, the above equation is:

$$\frac{u_{i,j}^{m-1} - 2u_{i,j}^{m} + u_{i,j}^{m+1}}{\Delta t^2} = \frac{u_{i-1,j}^{m} + u_{i+1,j}^{m} + u_{i,j-1}^{m} + u_{i,j+1}^{m} - 4u_{i,j}^{m}}{\Delta x^2} \tag{2}$$

Then isolate the term that marches in time:

$$u_{i,j}^{m+1} = 2u_{i,j}^{m} - u_{i,j}^{m-1} + \Delta t^2 \left(\frac{u_{i-1,j}^{m} + u_{i+1,j}^{m} + u_{i,j-1}^{m} + u_{i,j+1}^{m} - 4u_{i,j}^{m}}{\Delta x^2}\right) \tag{3}$$

I then simulate this system using MATLAB(wave2Dsolver.m provided in Github page is the code for this simulation). Run the simulation, you can visualize the simulation of this 2D wave equation. Here I have some captures of the simulation of the wave equations:



Then I get the analytic solution of this 2D wave equation for the log-log plot of the error vs. the grid spacing. Using separate variable method can solve this equation. The final analytical solution is:

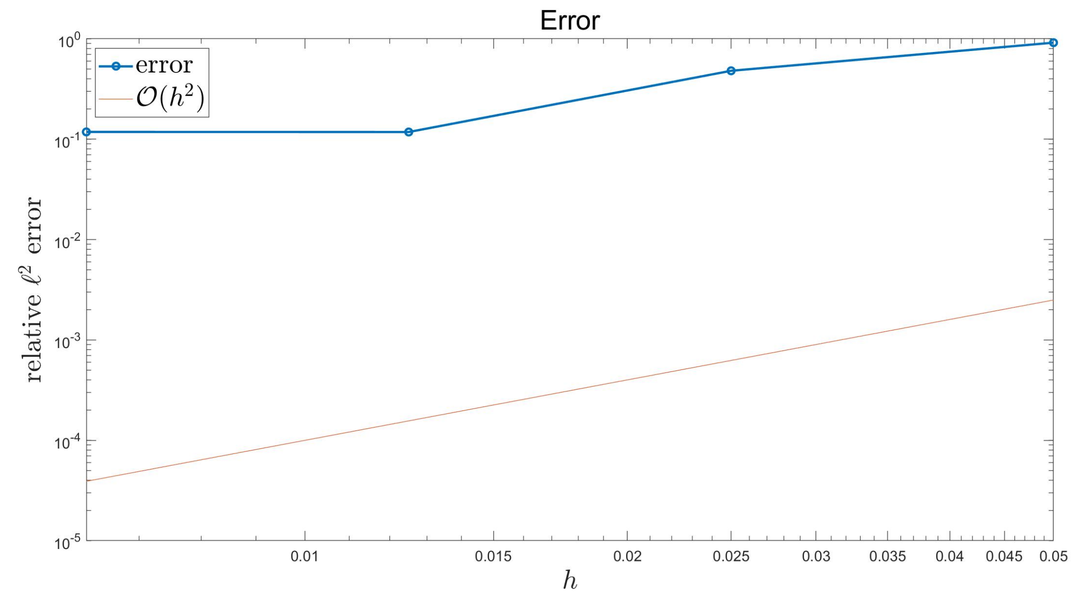
$$u(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \sin(m\pi x) \sin(n\pi y) B_{mn} \sin(\lambda_{mn} t) \tag{4}$$

where,

$$\lambda_{mn} = \sqrt{m^2 + n^2}\pi \tag{5}$$

$$B_{mn} = \frac{4}{\lambda_{mn}} \int_0^1 \int_0^1 e^{-400(x-0.5)^2 - 400(y-0.5)^2} \sin(m\pi x)\sin(n\pi y)\, dy\, dx \tag{6}$$

Two ways to calculate the error, one is to compute the uniform norm of the difference between analytic solution and the simulation solution, and the other one is to use a very fine grid scheme to represent analytic solution. I think since the analytic solution here is too complicated, I decided to use very fine grid scheme instead. And the log-log plot is:



From the plot we can tell that the method is second-order accurate.

2.2 Question (b)

The 3-point rule for y'' is:

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{(\Delta t)^2} = \lambda y^n \tag{7}$$

Only keep $n+1$ term at RHS, we get:

$$y^{n+1} = 2y^n - y^{n-1} + \lambda(\Delta t)^2 y^n \tag{8}$$

$$y^{n+1} = (2 + \lambda(\Delta t)^2)y^n - y^{n-1} \tag{9}$$

$$y^{n+1} - (2 + \lambda(\Delta t)^2)y^n + y^{n-1} = 0 \tag{10}$$

Since we have $y^n = \rho^n$, we then can generate

$$\rho^2 - (2 + \lambda(\Delta t)^2)\rho + 1 = 0 \tag{11}$$

To find stability region, we need

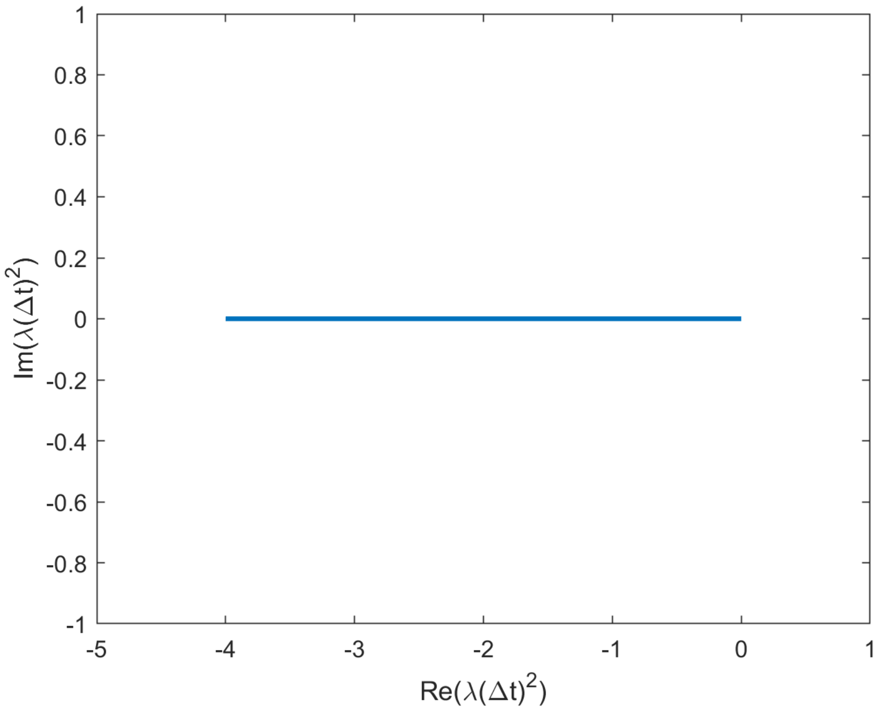
$$|\rho| \leq 1 \tag{12}$$

Now, set b = $\lambda(\Delta t)^2$, then to satisfy equation (12), we need

$$|\frac{2 + b \pm \sqrt{(2+b)^2 - 4}}{2}| \leq 1 \tag{13}$$

Therefore, the stability region is:

$$-4 \leq Re(\lambda(\Delta t)^2) \leq 0 \tag{14}$$



2.3 Question (c) - Question (e)

Will do them later

3 Probem D

I worked on this problem with Diya Yang.

For example, we consider a second order ODE:

$$u''(t) + au'(t) + bu(t) + c(t) = 0 \tag{15}$$

with boundary conditions $u(0) = u_0$ and $u'(0) = u'_0$. This is a linear second order ODE, in other words, a and b here are both constants. The numerical scheme for this problem is:

$$\frac{U_{n+1} - 2U_n + U_{n-1}}{(\Delta t)^2} + \frac{a}{2\Delta t}(U_{n+1} - U_{n-1}) + bU_n + c_n = 0 \tag{16}$$

We then keep all U_{n+1} terms on the LHS and the other terms on the RHS and we get:

$$\frac{2 + a\Delta t}{2(\Delta t)^2}U_{n+1} = (\frac{2}{\Delta t}^2 - b)U_n + \frac{a\Delta t - 2}{2(\Delta t)^2}U_{n-1} - c_n \tag{17}$$

$$\frac{2 + a\Delta t}{2(\Delta t)^2}U_{n+1} = (\frac{2}{\Delta t}^2 - b)U_n + \frac{a\Delta t - 2}{2(\Delta t)^2}U_{n-1} - c_n \tag{18}$$

We then divided $\frac{2+a\Delta t}{2(\Delta t)^2}$ for both side and set the coefficients as function $A(\Delta t)$, $B(\Delta t)$, and $C_n(\Delta t)$.This gives equation (18) as:

$$U_{n+1} = A(\Delta t)U_n + B(\Delta t)U_{n-1} + C_n(\Delta t) \tag{19}$$

The exact solution, denoted as u is:

$$u_{n+1} = Au_n + B)u_{n-1} + C_n + \Delta t\tau_n \tag{20}$$

where τ_n represents the local truncation error (LTE). The global error is denoted as:

$$E_n = U_n - u_n \tag{21}$$

And the numerical scheme for it is:

$$E_{n+1} = AE_n + BE_{n-1} - \Delta t \tau_n \tag{22}$$

Expand equation (22) in the more general term considering boundary conditions

$$E_N = \sum_{k=0}^{\frac{N-2}{2}} \binom{N-k-2}{k} A^{N-2k-2} B^{k+1} E_0 + \sum_{k=0}^{\frac{N-1}{2}} \binom{N-k-1}{k} A^{N-2k-1} B^k E_1 \tag{23}$$

$$+\Delta t \sum_{n=0}^{N-1} \sum_{k=0}^{\frac{N-n-1}{2}} \binom{N-k-n-1}{k} A^{N-n-2k-1} B^k \tau_n \tag{24}$$

Due to fixed BC, we have $E_0 = 0$ and $E_1 \to 0$ as $\Delta t \to 0$. And Because of consistency, we have $\tau_n \to 0$ as well. Due to Weak Stability $\Rightarrow$ Convergence. In this problem, Weak stability states that $\forall T > 0$, $\exists C_T > 0$ s.t. $||\sum_{k=0}^{\frac{N-1}{2}} \binom{N-k-1}{k} A^{N-2k-1} B^k|| \leq C_T$, $\forall \Delta t$, s.t., $n\Delta t \leq T$. Similarly, $||\sum_{k=0}^{\frac{N-n-1}{2}} \binom{N-k-n-1}{k} A^{N-n-2k-1} B^k||$ are also bounded by C'_T. Therefore, we have

$$E_N \leq C_T E_1 + \Delta t N C'_T \max_n \tau_n \tag{25}$$

where, as $\Delta t \to 0$

$$C_T E_1 + \Delta t N C'_T \max_n \tau_n \to 0 \tag{26}$$

Therefore, this system weak stable and is convergent.

Whereas to prove Not Weakly Stable $\Rightarrow$ Not Convergent,
If this system is not weakly stable, then there exists $T > 0$ s.t. for any $C_T > 0$ there exists Δt with $n\Delta t \leq T$ s.t. $||\sum_{k=0}^{\frac{N-1}{2}} \binom{N-k-1}{k} A^{N-2k-1} B^k|| \geq C_T$. Now we take $C_T > \frac{1}{||E_1||}$, then $\sum_{k=0}^{\frac{N-1}{2}} \binom{N-k-1}{k} A^{N-2k-1} B^k$ will not converge to 0. Similarly for terms has τ_n, $\sum_{k=0}^{\frac{N-n-1}{2}} \binom{N-k-n-1}{k} A^{N-n-2k-1} B^k$, and these terms cannot be cancelled. Therefore, this system is not convergent if it is not weakly stable.

Overall, above is proof of one extension of the Lax equivalence theorem to the case of linear ODE with two time derivatives.

Hw 2. MATH 714. KANGOZ FU

Problem A.

(a).

$\because$ $v \in \text{span}\{w_1, w_2, \cdots, w_n\}$. $\therefore$ we can write v as:

$v = a_1 w_1 + a_2 w_2 + \cdots + a_n w_n$, where $\{a_j\}$ is a collection of coefficients.

$= \sum_{j=1}^{n} a_j w_j$

now, to find a_j, since $\langle w_i, w_j \rangle = 0$ where $i \neq j$. we can have:

$v \cdot w_j = (a_1 w_1 + a_2 w_2 + \cdots + a_n w_n) \cdot w_j = \overset{0}{\cancel{a_1 w_1 \cdot w_j}} + \overset{0}{\cancel{a_2 w_2 \cdot w_j}} + \overset{0}{\cancel{\cdots}} + a_j w_j \cdot w_j + \overset{0}{\cancel{a_n w_n \cdot w_j}}$

$= a_j \|w_j\|^2$

$\therefore$ $a_j = \frac{v \cdot w_j}{\|w_j\|^2} = \frac{\langle v, w_j \rangle}{\|w_j\|^2}$ $\therefore$ overall, we have $\boxed{v = \sum_{j=1}^{n} \frac{\langle v, w_j \rangle}{\|w_j\|^2} w_j}$

(b) Note, two vectors $d_{(i)}$ & $d_{(j)}$ are A-orthogonal if $d_{(i)}^T A d_{(j)} = 0$

i. Because for CG, it will definitely converge at step N, therefore the optimal result, x^*, might be "<" N.

ii. Prove by induction:

$\langle p_1, p_0 \rangle_A = \langle r_1 - \frac{\langle r_1, p_0 \rangle_A}{\|p_0\|_A^2} p_0 , p_0 \rangle_A = \langle r_1 - \frac{\langle r_1, r_0 \rangle_A}{\|r_0\|_A^2} r_0 , r_0 \rangle_A = \langle r_1, r_0 \rangle_A - \langle \frac{\langle r_1, r_0 \rangle_A}{\|r_0\|_A^2} r_0, r_0 \rangle_A$

$= \langle r_1, r_0 \rangle_A - \left[\frac{r_1^T A r_0}{\|r_0\|_A^2} r_0\right]^T A r_0 = \langle r_1, r_0 \rangle_A - \left(\frac{r_1^T A r_0}{\cancel{\|r_0\|_A^2}}\right) \underbrace{\cancel{r_0^T A r_0}}_{\|r_0\|_A^2} = \langle r_1, r_0 \rangle_A - \underbrace{r_1^T A r_0}_{\langle r_1, r_0 \rangle_A} = 0$

$\langle p_2, p_0 \rangle_A = \langle r_2 - \frac{\langle r_2, p_0 \rangle_A}{\|p_0\|_A^2} p_0 - \frac{\langle r_2, p_1 \rangle}{\|p_1\|_A^2} p_1 , p_0 \rangle_A = (A p_0) \cdot \left(r_2 - \frac{r_2 \cdot A p_0}{\|p_0\|_A^2} p_0 - \frac{r_2 \cdot A p_1}{\|p_1\|_A^2} p_1 \right)$

$= (A p_0)^T r_2 - (A p_0)^T \left(\frac{r_2 \cdot A p_0}{\|p_0\|_A^2}\right) p_0 - (A p_0)^T \left(\frac{r_2 \cdot A p_1}{\|p_1\|_A^2} p_1\right)$

$= (A p_0)^T r_2 - \cancel{p_0^T A^T p_0} \left(\frac{r_2 \cdot A p_0}{\cancel{\|p_0\|_A^2}}\right) - \overset{0}{\cancel{(A p_0)^T p_1}} \left(\frac{r_2 \cdot A p_1}{\|p_1\|_A^2}\right)$ from $n=1$, we already have $\langle p_1, p_0 \rangle_A = 0$ $\therefore$ the last term is 0

$= (A p_0)^T r_2 - r_2 \cdot A p_0$ ($\because$ vector $a \cdot b = b \cdot a$)

$= 0$

$\langle p_2, p_1 \rangle_A = \left(r_2 - \frac{r_2 \cdot A p_0}{\|p_0\|_A^2} p_0 - \frac{r_2 \cdot A p_1}{\|p_1\|_A^2} p_1 \right) \cdot A p_1 = (r_2 \cdot A p_1) - \underset{\because p_0 \cdot A p_1 = 0}{0} - (r_2 \cdot A p_1) = 0$

$\vdots$

therefore, we can continue this process and compute out that

$\langle p_3, p_2 \rangle_A = \langle p_3, p_1 \rangle_A = \langle p_3, p_1 \rangle_A = 0$

$\therefore$ now, by mathematical induction, assume, we have P_{K-1} is A-conjugate to $P_0, \cdots P_{K-2}$, Let $P_j \in \{P_0, \cdots, P_{K-1}\}$. we need to show that $\langle P_K, P_j \rangle_A = 0$

$$\langle P_K, P_j \rangle_A = P_K \cdot AP_j = \left(r_K - \sum_{i=1}^{K-1} \frac{\langle r_K, P_i \rangle_A}{\|P_i\|_A^2} P_i\right) \cdot AP_j = r_K \cdot AP_j - \sum_{i=1}^{K-1} \frac{\langle r_K, P_i \rangle_A P_i}{\|P_i\|_A^2} \cdot AP_j$$

$$= r_K \cdot AP_j - 0 - 0 - \cdots - \frac{\langle r_K, P_j \rangle_A P_j}{\|P_j\|_A^2} \cdot AP_j = \langle r_K, P_j \rangle_A - \frac{\langle r_K, P_j \rangle_A (\cancel{P_j \cdot AP_j})}{\cancel{\|P_j\|_A^2}} = \underline{\underline{0}}$$

$\therefore$ now, we proved that $\langle P_n, P_j \rangle_A = 0$, for $0 \leq j < n \leq N-1$ ∎

(c). i.) $\langle Av, w \rangle = (Av)^T w = v^T A^T w \overset{\because A = A^T}{=} v^T A w$, $\because A$ is symmetric, $\therefore A$ is diagonalizable and can be decomposed in the form: $A = PDP^{-1}$, where $P = [\phi_1 \ \phi_2 \cdots \phi_n]$; $D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$,

since $\langle \phi_n, \phi_j \rangle = \delta_{nj}$, $\therefore P$ is an orthogonal matrix and $P^{-1} = P^T$ $\therefore PDP^T = \sum_{n=1}^{N} \lambda_n \phi_n \phi_n^T$

$$\therefore \langle Av, w \rangle = v^T A w = v^T PDP^{-1} w = v^T PDP^T w = v^T \left(\sum_{n=1}^{N} \lambda_n \phi_n \phi_n^T\right) w = \sum_{n=1}^{N} v^T \lambda_n \phi_n \phi_n^T w$$

$$= \sum_{n=1}^{N} \lambda_n \langle v, \phi_n \rangle \langle \phi_n, w \rangle$$ ∎

ii) $\because A$ is symmetric positive definite $\therefore A$ can be decomposed by $A = S^T S$

$\therefore \quad S^T S \phi_n = \lambda_n \phi_n$

$\therefore \quad \phi_n^T S^T S \phi_n = \phi_n^T \lambda_n \phi_n = \lambda_n \phi_n^T \phi_n$, $\quad LHS = (S\phi_n)^T (S\phi_n) = \|S\phi_n\|^2$, $\quad RHS = \lambda_n \|\phi_n\|^2$

$\therefore \quad \lambda_n = \dfrac{\|S\phi_n\|^2}{\|\phi_n\|^2}$ $\quad \because \lambda_n \neq 0$ because $S^T S = A$ is invertable

$\therefore$ overall $\lambda_n > 0$, for $1 \leq n \leq N$ ∎

iii). $\langle Av, v \rangle = (Av)^T v = v^T A^T v = v^T A v = \sum_{n=1}^{N} \lambda_n \langle v, \phi_n \rangle \langle \phi_n, v \rangle$ from i) $= v^T \left(\sum_{n=1}^{N} \lambda_n \phi_n \phi_n^T\right) v$

$\lambda_1 \|v\|^2 = \lambda_1 v^T v = v^T \lambda_1 v = v^T = v^T \lambda_1 I v^T = v^T \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_1 & \\ 0 & & \ddots \ \lambda_1 \end{bmatrix} v$

$$= v^T \left(\sum_{n=1}^{N} \lambda_1 \phi_n \phi_n^T\right) v$$

$$\lambda_N \|v\|^2 = v^T \left(\sum_{n=1}^{N} \lambda_N \phi_n \phi_n^T\right) v$$

$\because \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$

$\therefore \underset{\sim}{v}^T \left(\sum_{n=1}^{N} \lambda_1 \underset{\sim}{\phi_n} \underset{\sim}{\phi_n}^T\right) \underset{\sim}{v} \le \underset{\sim}{v}^T \left(\sum_{n=1}^{N} \lambda_n \phi_n \phi_n^T\right) \underset{\sim}{v} \le \underset{\sim}{v}^T \left(\sum_{n=1}^{N} \lambda_N \phi_n \phi_n^T\right) \underset{\sim}{v}$

$\therefore$ overall.

$$\underbrace{\lambda_1 \|v\|^2 \le \langle Av, v\rangle \le \lambda_N \|v\|^2}$$ ■

iv. write $\underset{\sim}{v} = c_1 \underset{\sim}{\phi_1} + c_2 \underset{\sim}{\phi_2} + \cdots + c_n \underset{\sim}{\phi_n}$

$\because \|A \underset{\sim}{\phi_n}\| = \|\lambda_n \underset{\sim}{\phi_n}\| = \lambda_n$

$\|A\underset{\sim}{v}\| = \|c_1 \lambda_1 \phi_1 + \cdots + c_N \lambda_N \phi_N\|$

$$\therefore \frac{\|A\underset{\sim}{v}\|^2}{\|\underset{\sim}{v}\|^2} = \frac{\|c_1\lambda_1\phi_1 + \cdots + c_N\lambda_N\phi_N\|^2}{c_1^2 + \cdots + c_n^2} = \frac{c_1^2\lambda_1^2 + c_2^2\lambda_2^2 + \cdots + c_N^2\lambda_N^2}{c_1^2 + \cdots + c_n^2} \le \lambda_N^2 \frac{\cancel{c_1^2 + \cdots + c_N^2}}{\cancel{c_1^2 + \cdots + c_N^2}}$$

$\therefore \|A\underset{\sim}{v}\|^2 \le \lambda_N^2 \|\underset{\sim}{v}\|^2 \quad \therefore \boxed{\|A\underset{\sim}{v}\| \le \lambda_N \|\underset{\sim}{v}\|}$ ■

(d). $p_{n+1} = r_{n+1} + \beta_n p_n = (r_n - \alpha_n w_n) + \beta_n p_n = (r_n - \alpha_n A p_n) + \beta_n p_n$

$= \beta_n p_n - \alpha_n A p_n + \underset{\sim}{r_n} = \beta_n p_n - \alpha_n A p_n + (\underbrace{p_n - \beta_{n-1} p_{n-1}}_{r_n}) = \underbrace{\{(\beta_n + 1) p_n - \alpha_n A p_n - \beta_{n-1} p_{n-1}\}}$ ■

(e). Because $A \in \mathbb{R}^{n\times n}$ is non-singular, $\therefore$ it is invertible and we can use Cayley-Hamilton, which states that: $\underline{p(A) = A^N + c_{n-1}A^{N-1} + \cdots + c_1 A_1 + (-1)^N \det(A) I_n = 0}$. ①

the characteristic polynomial of A: $p(\lambda) = \lambda^N + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$

which is in the same format of

$\therefore$ Based on ①, we have:

$$A^N = \underbrace{-c_{n-1}A^{N-1} - \cdots - c_1 A_1 - (-1)^N \det(A) I}$$

$\therefore$ we have A^N as a linear combination

(f)

i. $Au = f$; $u_{n+1} = u_n + \alpha(f - Au_n)$

$e_n = u_n - u$ ∴ $e_{n+1} = u_{n+1} - u = u_n + \alpha(f - Au_n) - u$

∴ $e_{n+1} = (u_n - u) + \alpha(f - Au_n) = e_n + \alpha(f - Au_n) = e_n + \alpha(Au - Au_n)$

$= e_n - \alpha A(u_n - u) = e_n - \alpha A e_n = (I - \alpha A)e_n$

∴ $e_{n+1} = (I - \alpha A)e_n$ ∎

ii) we have:

$\|e_{n+1}\| = \|(I-\alpha A)e_n\| \le \|I - \alpha A\|\,\|e_n\|$. For $\|I - \alpha A\|$, we have: $\|I - \alpha A\| \le |\|I\| - \|\alpha A\||$

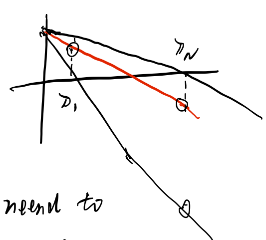
Spectral radius: $\rho(A) = \max\{|\lambda_1|, \dots |\lambda_n|\}$ ∴ $\|\alpha A\| \ge \alpha\rho(A)$ ∴ $\|I - \alpha A\| \le |1 - \rho(A)|$

∴ $\|e_{n+1}\| \le |1 - \rho(A)|\,\|e_n\|$, where $|1-\rho(A)| = \max_{1 \le j \le N} |1 - \alpha\lambda_j| = \rho$

∴ $\|e_{n+1}\| \le \rho\|e_n\|$ ∎ ∴ $\|e_n\| \le \rho\|e_{n-1}\|$ ---- ⇒ ∴ $\|e_{n+1}\| \le \rho^n\|e_0\|$

∴ if $\rho < 1$, then we have: $\rho^n \to 0$ as $n \to \infty$ ∴ $\|e_{n+1}\| \to 0$ which means $\boxed{\lim_{n\to\infty} u_n = u}$

iii). $\rho = \max_{1 \le j \le N} |\overset{P(\lambda)}{1 - \alpha\lambda_j}|$,



set $P(\lambda) = 1 - \alpha\lambda$,

∴ to have minimum value, we need to have: $|P(\lambda_1)| = |P(\lambda_N)|$ (shown from above graph, we can visualize if "≠" then either λ_1 or λ_N, which gives the max $|1-\alpha\lambda_j|$ will be larger. ∴ To minimize it, we need to have: $P(\lambda_1) = -P(\lambda_N)$

∴ then we have: $1 - \alpha\lambda_1 = -(1 - \alpha\lambda_N)$ ∴ $1 - \alpha\lambda_1 = -1 + \alpha\lambda_N$ ∴ $2 = \alpha(\lambda_1 + \lambda_N)$ ∴ $\alpha = \frac{2}{\lambda_1 + \lambda_N}$

∴ If $\alpha = \frac{2}{\lambda_1 + \lambda_N}$, we have $\rho_{min} = \min \max_{1 \le j \le N} |1 - \alpha\lambda_j| = 1 - \frac{2}{\lambda_1 + \lambda_N}\lambda_1 = \frac{\lambda_1 + \lambda_N - 2\lambda_1}{\lambda_1 + \lambda_N}$

$= \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} = \frac{(\lambda_N - \lambda_1)\cdot\frac{1}{\lambda_1}}{(\lambda_N + \lambda_1)\cdot\frac{1}{\lambda_1}} = \boxed{\frac{k-1}{k+1}}$ ①, where $k = \frac{\lambda_N}{\lambda_1}$ ∎

? iv). since $0 < c \le \lambda_1 \le \lambda_N \le C < \infty$

same as iii, the sub-optimal choice is $1 - \alpha\tilde{\lambda}_1 = -1(1 - \alpha\tilde{\lambda}_n) = 1 - \alpha c = -1(1 - \alpha C) \to \alpha = \frac{2}{c + C}$

then $\hat{\rho} = 1 - \frac{2}{c + C}c = \frac{C - c}{C + c}$ ∴ $\hat{\rho} = \frac{C - c}{C + c} = \frac{\frac{C}{c} - \frac{c}{c}}{\frac{C}{c} - \frac{c}{c}} = \frac{k' - 1}{k' + 1} < 1$. From (iii), we have

$\rho = \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1}$ from $c \le \lambda_1 \le \lambda_N \le C$ we have: $c\lambda_N \le \lambda_1\lambda_N$, $c\lambda_N \le \lambda_1 C$

∴ $2c\lambda_N \le \lambda_1\lambda_N + \lambda_1 C \to -\lambda_1 C + c\lambda_N \le \lambda_1\lambda_N - c\lambda_N$

∴ $C\lambda_N - c\lambda_1 - \lambda_1 C + c\lambda_N \le C\lambda_N - c\lambda_1 + \lambda_1\lambda_N - c\lambda_N$

↓

$\therefore (C+c)(\lambda_N-\lambda_1) \le (C-c)(\lambda_1+\lambda_N)$ $\quad \therefore \dfrac{\lambda_N-\lambda_1}{\lambda_1+\lambda_N} \le \dfrac{C-c}{C+c}$ $\quad \therefore$ $\boxed{\text{we have: } \rho \le \dfrac{C-c}{C+c} = \dfrac{k'-1}{k'+1} < 1, \ k' = \dfrac{C}{c}}$

(g) From notes, we have:

$$q_n = \frac{r_n}{\|r_n\|},\ 0 \le n \le n^*-1, \quad p_0 = r_0 = f - Au_0, \quad w_n = Ap_n, \quad r_{n+1} = r_n - \alpha_n w_n, \quad p_n = r_n + \beta_{n-1} p_{n-1}, \quad \beta_n = \frac{r_{n+1}^T r_{n+1}}{r_n^T r_n}$$

$\therefore$

(i) $r_1 = r_0 - \alpha_0 w_0 = r_0 - \alpha_0 A p_0 = \underline{r_0 - \alpha A r_0}$ $\square$

(ii) $r_{n+1} = r_n - \alpha_n w_n = r_n - \alpha_n A p_n$ $\quad \therefore A p_n = \dfrac{r_{n+1} - r_n}{-\alpha_n}$

$$= r_n - \alpha_n A(r_n + \beta_{n-1} p_{n-1}) = r_n - \alpha_n A r_n - \alpha_n \beta_{n-1} A p_{n-1} = r_n - \alpha_n A r_n - \alpha_n \beta_{n-1} \frac{r_{n-1} - r_n}{\alpha_{n-1}}$$

$\therefore$ $\boxed{r_{n+1} = r_n - \alpha A r_n + \dfrac{\alpha_n \beta_{n-1}}{\alpha_{n-1}}(r_n - r_{n-1})}$ for $1 \le n \le n^*-1$ $\square$

(iii) Now, we have new variable: $\gamma_0 = \dfrac{1}{\alpha_0}$, $\gamma_n = \dfrac{1}{\alpha_n} + \dfrac{\beta_{n-1}}{\alpha_{n-1}}$ for $1 \le n \le n^*-1$, $\delta_n = \dfrac{\sqrt{\beta_n}}{\alpha_n}$

in (i), $A r_0 = \dfrac{r_0 - r_1}{\alpha_0}$ $\quad \therefore A q_0 = \dfrac{A r_0}{\|r_0\|} = \dfrac{r_0 - r_1}{\alpha_0 \|r_0\|} = \dfrac{1}{\alpha_0}\dfrac{r_0}{\|r_0\|} - \dfrac{1}{\alpha_0}\dfrac{r_1}{\|r_0\|} = \dfrac{1}{\alpha_0}\dfrac{r_0}{\|r_0\|} - \dfrac{\sqrt{\beta_0}}{\alpha_0}\dfrac{r_1}{\|r_1\|}$

$= \boxed{\gamma_0 q_0 - \delta_0 q_1}$ $\quad \therefore$ $\boxed{A q_0 = \gamma_0 q_0 - \delta_0 q_1}$

in (ii),

$$r_{n+1} = r_n - \alpha_n A r_n + \frac{\alpha_n \beta_{n-1}}{\alpha_{n-1}}(r_n - r_{n-1}) \quad \therefore A r_n = \frac{r_n - r_{n+1}}{\alpha_n} + \frac{\beta_{n-1}}{\alpha_{n-1}}(r_n - r_{n-1})$$

$$\therefore A q_n = \frac{A r_n}{\|r_n\|} = \frac{1}{\alpha_n}\frac{r_n}{\|r_n\|} - \frac{1}{\alpha_n}\frac{r_{n+1}}{\|r_n\|} + \frac{\beta_{n-1} r_n}{\alpha_{n-1}\|r_n\|} - \frac{\beta_{n-1}}{\alpha_{n-1}}\frac{r_{n-1}}{\|r_n\|}$$

$\therefore A q_n = \left(\dfrac{1}{\alpha_n} + \dfrac{\beta_{n-1}}{\alpha_{n-1}}\right)\dfrac{r_n}{\|r_n\|} - \dfrac{\sqrt{\beta_n}}{\alpha_n}\dfrac{r_{n+1}}{\|r_{n+1}\|} - \dfrac{\sqrt{\beta_{n-1}}}{\alpha_{n-1}}\dfrac{r_{n-1}}{\|r_{n-1}\|} = \boxed{\gamma_n q_n - \delta_n q_{n+1} - \delta_{n-1} q_{n-1}}$ for $1 \le n \le n^*-1$

(iv). $Q_n = [q_0 \cdots q_{n-1}] \in \mathbb{R}^{N \times n}$

$$T_n = \begin{bmatrix} \gamma_0 & -\delta_0 & & & \\ -\delta_0 & \gamma_1 & -\delta_1 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & -\delta_{n-3} & \gamma_{n-2} & -\delta_{n-2} \\ & & & -\delta_{n-2} & \gamma_{n-1} \end{bmatrix} \qquad e_n = [0\ 0 \cdots 1]^T \in \mathbb{R}_n$$

$N \times n \qquad n \times n \rightarrow N \times n$

$\therefore Q_n T_n = [(q_0\gamma_0 - q_1\delta_0), (-\delta_0 q_0 + q_1\gamma_1 - q_2\delta_1) \cdots\cdots, (-q_{n-3}\delta_{n-3} + \gamma_{n-2} q_{n-2} - \delta_{n-2} q_{n-1}), (-q_{n-2}\delta_{n-2} + q_{n-1}\gamma_{n-1})]$

$\delta_{n-1} q_n e_n^T = [\ 0,\quad 0 \quad \cdots\cdots \quad 0,\quad -\delta_{n-1} q_n\ ]$

$\therefore$ $\underbrace{Q_n T_n - \delta_{n-1} q_n e_n^T}_{\text{(From iii)}} = [\ \underbrace{q_0 r_0 - q_1 \delta_0}_{\leftarrow Aq_0} \quad \underbrace{(-\delta_0 q_0 + q_1 r_1 - q_2 \delta_1)}_{Aq_1} \quad \cdots\cdots \quad \underbrace{(-q_{n-3}\delta_{n-3} + r_{n-2} q_{n-2} - \delta_{n-2} q_{n-1})}_{Aq_{n-2}}, \quad \underbrace{(-q_{n-2}\delta_{n-2} + q_{n-1} r_{n-1} - \delta_{n-1} q_n)}_{Aq_{n-1}}\ ]$

$= \underbrace{AQ_n}$

(v). From (iv), we can get:

$$Q_n^T A Q_n = \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_{n-1} \end{bmatrix} [(q_0 r_0 - q_1 \delta_0) \quad (-\delta_0 q_0 + q_1 r_1 - q_2 \delta_1) \quad \cdots\cdots, \quad (-q_{n-3}\delta_{n-3} + r_{n-2} q_{n-2} - \delta_{n-2} q_{n-1}), \quad (-q_{n-2}\delta_{n-2} + q_{n-1} r_{n-1})]$$

notice, $\because$ Q is orthogonal $\therefore$ $q_i q_j = \delta_{ij}$

$\therefore$ For example: $q_0(q_0 r_0 - q_1 \delta_0) = \overset{1}{q_0 q_0} r_0 - \overset{0}{\cancel{q_0 q_1}} \delta_0 = r_0$,

$q_0(-\delta_0 q_0 + q_1 r_1 - q_2 \delta_1) = -\delta_0$,

$q_1(q_0 r_0 - q_1 \delta_0) = -\delta_0$, $\quad q_1(-\delta_0 q_0 + q_1 r_1 - q_2 \delta_1) = r_1$,

$\vdots$

$q_{n-1}(-q_{n-2}\delta_{n-2} + q_{n-1} r_{n-1}) = r_{n-1}$

$\therefore$ we can have: $\boxed{Q_n^T A Q_n = T_n}$ $\square$

Problem B and C will on Latex