THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
MATH3310 2022-2023
Homework Assignment 5 Suggested Solution

1. Consider the linear system $A\boldsymbol{x} = \boldsymbol{k}$, where

$$A = \begin{pmatrix} 1 & 0 & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 1 & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & 0 & 1 \end{pmatrix} \text{ and } \boldsymbol{k} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Let $\boldsymbol{x}^* = (1,1,1,1)^T$ be the solution of the linear system. Suppose $\{\boldsymbol{x}^{(m)}\}_{m=1}^{\infty}$ and $\{\boldsymbol{y}^{(m)}\}_{m=1}^{\infty}$ are the sequences of vectors obtained by the Jacobi method and Gauss-Seidel method respectively to solve the linear system with initialization $\boldsymbol{x}^{(0)} = \boldsymbol{y}^{(0)} = (0,0,0,0)^T$. Let $\boldsymbol{e}_J^{(m)} := \boldsymbol{x}^{(m)} - \boldsymbol{x}^*$ and $\boldsymbol{e}_{GS}^{(m)} := \boldsymbol{y}^{(m)} - \boldsymbol{x}^*$ be the error vectors at the m-th iteration for the Jacobi and Gauss-Seidel method respectively.

(a) Show that: $\boldsymbol{e}_J^{(m)} = -2^{-m}(1,1,1,1)^T$ for $m \geq 1$.

(b) Show that: $\boldsymbol{e}_{GS}^{(m)} = -4^{-m}(2,2,1,1)^T$ for $m \geq 1$.

(c) Show that $\|\boldsymbol{e}_{GS}^{(m)}\|_2 < \|\boldsymbol{e}_J^{(m)}\|_2$ for $m \geq 1$. Hence, the Gauss-Seidel method converges faster than the Jacobi method.

Solution:

(a) For Jacobi method, we have

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; P = \begin{pmatrix} 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 \\ 1/4 & 1/4 & 0 & 0 \\ 1/4 & 1/4 & 0 & 0 \end{pmatrix}$$

so

$$\begin{aligned} e_J^{(m)} &= x^{(m)} - x^* \\ &= (N^{-1}P)^m e_J^{(0)} \\ &= P^m e_J^{(0)} \\ &= -\frac{1}{2^m}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

(b) For Gauss Seidel method, we have

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/4 & -1/4 & 1 & 0 \\ -1/4 & -1/4 & 0 & 1 \end{pmatrix}; P = \begin{pmatrix} 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

so

$$\begin{aligned} e_{GS}^{(m)} &= x^{(m)} - x^* \\ &= (N^{-1}P)^m e_{GS}^{(0)} \\ &= \begin{pmatrix} 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 1/4 & 1/4 \\ 0 & 0 & 1/8 & 1/8 \\ 0 & 0 & 1/8 & 1/8 \end{pmatrix}^m e_{GS}^{(0)} \\ &= \begin{pmatrix} 0 & 0 & 1/4^m & 1/4^m \\ 0 & 0 & 1/4^m & 1/4^m \\ 0 & 0 & 1/(2\cdot 4^m) & 1/(2\cdot 4^m) \\ 0 & 0 & 1/(2\cdot 4^m) & 1/(2\cdot 4^m) \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \\ &= -\frac{1}{4^m} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

(c) From the definition, we have $\|e_{GS}^{(m)}\|_2^2 = \frac{4+4+1+1}{4^{2m}} = \frac{10}{4^{2m}}$;
$\|e_J^{(m)}\|_2^2 = \frac{1+1+1+1}{2^{2m}} = \frac{4\cdot 4^{2m}}{4^{2m}}$
Clearly, we have $10 < 4 \cdot 4^{2m}$ for $m > 1$, hence $\|e_{GS}^{(m)}\|_2 < \|e_J^{(m)}\|_2$ for $m > 1$.

2. Consider the linear system $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite. Consider:

$$f(\boldsymbol{\eta}) = \frac{1}{2}\boldsymbol{\eta}^T A \boldsymbol{\eta} - \boldsymbol{b}^T \boldsymbol{\eta}$$

(a) For an iterative scheme $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}_k$, where $\boldsymbol{p}_k$ is a fixed direction, find α_k such that $f(\boldsymbol{x}^{(k+1)})$ is minimized.

(b) The conjugate gradient method is given by

$$\begin{aligned} \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}_k, \\ \boldsymbol{p}_{k+1} &= -\boldsymbol{r}_{k+1} - \beta_k \boldsymbol{p}_k, \\ \beta_k &= -\frac{\langle \boldsymbol{r}_{k+1}, \boldsymbol{p}_k \rangle_A}{\langle \boldsymbol{p}_k, \boldsymbol{p}_k \rangle_A} \end{aligned}$$

where α_k is in the form given by (a), $\boldsymbol{r}_k = A\boldsymbol{x}^{(k)} - \boldsymbol{b}$, $\boldsymbol{p}_0 = -\boldsymbol{r}_0$, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_A = \boldsymbol{x} \cdot A\boldsymbol{y}$, $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ is an arbitray initial guess. Provided that $\boldsymbol{r}_i \cdot \boldsymbol{r}_j = 0$ and $\langle \boldsymbol{p}_i, \boldsymbol{p}_j \rangle_A = 0$ for all $i \neq j$, show that $\beta_k = -\frac{\boldsymbol{r}_{k+1}\cdot \boldsymbol{r}_{k+1}}{\boldsymbol{r}_k \cdot \boldsymbol{r}_k}$.

Solution:

(a) Refer to Lecture 22.

(b) Note

$$\begin{aligned} \mathbf{r_{k+1}} &= A\mathbf{x_{k+1}} - \mathbf{b} \\ &= A(\mathbf{x_k} + \alpha_k \mathbf{p_k})\mathbf{b} \\ &= A\mathbf{x_k} - \mathbf{b} + \alpha_k A\mathbf{p_k} \\ &= \mathbf{r_k} + \alpha_k A\mathbf{p_k} \end{aligned}$$

So,

$$\begin{aligned}
\beta_{k+1} &= -\frac{\langle \mathbf{r_{k+1}}, \mathbf{p_k} \rangle_A}{\langle \mathbf{p_k}, \mathbf{p_k} \rangle_A} \\
&= -\frac{\mathbf{r_{k+1}}^T A \mathbf{p_k}}{-(\mathbf{r_k} + \beta_{k-1}\mathbf{p_{k-1}})^T A \mathbf{p_k}} \\
&= \frac{\mathbf{r_{k+1}}^T (\mathbf{r_{k+1}} - \mathbf{r_k})/\alpha_k}{\mathbf{r_k}^T A \mathbf{p_k} + \beta_{k-1}(\mathbf{p_{k-1}}^T A \mathbf{p_k})} \\
&= \frac{\mathbf{r_{k+1}}^T \mathbf{r_{k+1}}/\alpha_k}{\mathbf{r_k}^T (\mathbf{r_{k+1}} - \mathbf{r_k})/\alpha_k} \\
&= -\frac{\mathbf{r_{k+1}}^T \mathbf{r_{k+1}}}{\mathbf{r_k}^T \mathbf{r_k}}
\end{aligned}$$

3. Consider the gradient descent method for solving $A\boldsymbol{x} = \boldsymbol{b}$ with some $\alpha \in \mathbb{R}$ and A to be a symmetric positive definite matrix:

$$\begin{aligned}
\boldsymbol{x}^{k+1} &= \boldsymbol{x}^k + \alpha \boldsymbol{d}^k \\
\boldsymbol{d}^k &= -(A\boldsymbol{x}^k - \boldsymbol{b})
\end{aligned}$$

Prove that the method converges if and only if $\alpha < \frac{2}{\lambda_j}$ for all j where λ_j are the eigenvalues of A.

Hint: suppose $\boldsymbol{\eta}$ is the solution to $A\boldsymbol{x} = \boldsymbol{b}$, them η satisfies

$$\boldsymbol{\eta} = \boldsymbol{\eta} + \alpha(A\boldsymbol{\eta} - \boldsymbol{b}).$$

Using this equation, start with the error vector $\boldsymbol{e}^k = \boldsymbol{x}^k - \boldsymbol{\eta}$ to make some observation.

Solution:
Suppose η is a solution to $Ax = b$. Then,

$$\begin{aligned}
e^k &= x^k - \eta \\
&= x^{k-1} - \alpha(Ax^{k-1} - b) - \eta + \alpha(A\eta - b) \\
&= (I - \alpha A)x^{k-1} - (I - \alpha A)\eta \\
&= (I - \alpha A)e^{k-1}
\end{aligned}$$

Thus, like what we learnt in iterative methods, the method converges if and only if $\rho(I - \alpha A) < 1$. Suppose A has eigenvalues $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_{max}$. The the eigenvalues of the matrix $I - \alpha A$ are

$$1 - \alpha\lambda_1 > 1 - \alpha\lambda_2 > \cdots > 1 - \alpha\lambda_{max}$$

To make sure $\rho(I - \alpha A) < 1$, we have to make sure

$$\begin{aligned}
&|1 - \alpha\lambda_j| < 1 \text{ for all } j \\
&1 - \alpha\lambda_j < 1 \text{ and } 1 - \alpha\lambda_j > -1
\end{aligned}$$

Thus, it means that $\alpha\lambda_j < 2$ for all j. Thus, $\alpha < \frac{2}{\lambda_{max}}$.