PS #5 — Iterative methods

Due: 3/10/25, 11:59 PM *Instructor: Jonathan Lindbloom*

Problem 1. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be positive definite. For Richardson's method

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha(\mathbf{b} - \mathbf{A}\mathbf{x}_k), \quad k = 1, 2, \ldots, \tag{1}$$

show that the spectral radius $\rho(\mathbf{T})$ of the iteration matrix $\mathbf{T}$ (governing the convergence rate) is minimized with the choice

$$\alpha_{\text{opt}} = \frac{2}{\lambda_{\min}(\mathbf{A}) + \lambda_{\max}(\mathbf{A})}. \tag{2}$$

Furthermore, for the choice $\alpha = \alpha_{\text{opt}}$ find an expression for $\rho(\mathbf{T})$ in terms of the condition number of $\mathbf{A}$.

Problem 2. Let $\mathbf{A}$ be an SPD matrix, and let $\varphi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x}$. Give a justification for the formulas

$$\nabla\varphi(\mathbf{x}) = \begin{bmatrix} \frac{\partial\varphi}{\partial x_1} & \cdots & \frac{\partial\varphi}{\partial x_1} \end{bmatrix}^T = \mathbf{A}\mathbf{x} - \mathbf{b}, \qquad \mathbf{H}_\varphi = \begin{bmatrix} \frac{\partial^2\varphi}{\partial x_1^2} & \frac{\partial^2\varphi}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2\varphi}{\partial x_1 \partial x_n} \\ \frac{\partial^2\varphi}{\partial x_2 \partial x_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial^2\varphi}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2\varphi}{\partial x_n^2} \end{bmatrix} = \mathbf{A}. \tag{3}$$

If helpful, you may consider a simple 2×2 or 3×3 matrix $\mathbf{A}$.

Problem 3. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a set of linearly independent vectors in $\mathbb{R}^n$, and let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be SPD. Explain how a set of $\mathbf{A}$-conjugate vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ can be computed (i.e., $\mathbf{v}_i^T\mathbf{A}\mathbf{v}_j = 0$ if $i \neq j$).

Problem 4. In this problem we derive some properties of the CG method. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be SPD. Recall that the CG method begins by defining an initial condition $\mathbf{x}_0 \in \mathbb{R}^n$ as well as an initial direction $\mathbf{d}_0 = \mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$. Then, subsequence approximate solutions $\mathbf{x}_k$ are computed via the iteration

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{d}_k, \quad \alpha_k = \frac{\langle \mathbf{r}_k, \mathbf{d}_k \rangle}{\langle \mathbf{d}_k, \mathbf{d}_k \rangle_{\mathbf{A}}}, \tag{4}$$

$$\mathbf{d}_{k+1} = \mathbf{r}_{k+1} + \beta_k\mathbf{d}_k, \quad \beta_k = -\frac{\langle \mathbf{r}_{k+1}, \mathbf{d}_k \rangle_{\mathbf{A}}}{\langle \mathbf{d}_k, \mathbf{d}_k \rangle_{\mathbf{A}}}, \tag{5}$$

for $k = 0, 1, \ldots$, until convergence.

Part (a): Prove the two relations

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k\mathbf{A}\mathbf{d}_k, \tag{6}$$

$$\mathbf{r}_{k+1} = \mathbf{d}_{k+1} - \beta_k\mathbf{d}_k, \tag{7}$$

for $k \geq 0$.

Part (b): Show that each of

$$\langle \mathbf{r}_1, \mathbf{r}_0 \rangle = 0, \tag{8}$$

$$\langle \mathbf{r}_1, \mathbf{d}_0 \rangle = 0, \tag{9}$$

$$\langle \mathbf{d}_1, \mathbf{d}_0 \rangle_{\mathbf{A}} = 0, \tag{10}$$

is satisfied. *Bonus part (c):* Assume that each of

$$\langle \mathbf{r}_k, \mathbf{r}_j \rangle = 0, \tag{11}$$
$$\langle \mathbf{r}_k, \mathbf{d}_j \rangle = 0, \tag{12}$$
$$\langle \mathbf{d}_k, \mathbf{d}_j \rangle_{\mathbf{A}} = 0, \tag{13}$$

is satisfied for some $k \in \mathbb{N}$ and for all $j \leq k-1$. Show that this implies that each of

$$\langle \mathbf{r}_{k+1}, \mathbf{r}_j \rangle = 0, \tag{14}$$
$$\langle \mathbf{r}_{k+1}, \mathbf{d}_j \rangle = 0, \tag{15}$$
$$\langle \mathbf{d}_{k+1}, \mathbf{d}_j \rangle_{\mathbf{A}} = 0, \tag{16}$$

is satisfied for all $j \leq k$.

Part (d): Use these relations to explain why $\mathbf{d}_{k+1}$ is always a descent direction of $\varphi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x}$ at the point $\mathbf{x}_{k+1}$ (i.e., show that $\langle \mathbf{d}_{k+1}, \nabla\varphi(\mathbf{x}_{k+1}) \rangle < 0$).

Problem 5. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Furthermore, suppose that $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\mathbf{H}$ where $\mathbf{V}_{k+1} = [\mathbf{V}_k, \mathbf{v}_{k+1}] = [\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}]$ has orthonormal columns and $\mathbf{H} \in \mathbb{R}^{(k+1) \times k}$ is upper Hessenberg. Suppose that $\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|_2$. Show that

$$\operatorname*{arg\,min}_{\mathbf{x} \in \mathrm{col}(\mathbf{V}_k)} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 = \mathbf{V}_k \left(\operatorname*{arg\,min}_{\mathbf{z} \in \mathbb{R}^k} \|\mathbf{H}\mathbf{z} - \beta\mathbf{e}_1\|_2^2 \right) \tag{17}$$

where $\mathbf{e}_1 = (1, 0, \ldots, 0)^T \in \mathbb{R}^{k+1}$ and $\beta = \|\mathbf{b}\|_2$.