Study Questions 5 Numerical Linear Algebra Spring 2024

Problem 5.1

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite and define the A-norm using the A-inner product

$$\langle v_1, v_2\rangle_A = v_2^T A v_1$$

$$\|v\|_A^2 = \langle v, v\rangle_A.$$

Consider the linear system $Ax = b$ with solution $x_* = A^{-1}b$. Define the two functions from $\mathbb{R}^n$ to $\mathbb{R}$

$$E(x) = \|x - x_*\|_A^2, \quad f(x) = \frac{1}{2}x^T A x - b^T x$$

(5.1.a) Show that $E(x)$ and $f(x)$ have the same unique minimizer x_*.

(5.1.b) If $Ax = b$ is solved using the general descent method the stepsize α_k, used in $x_{k+1} = x_k + \alpha_k p_k$, is defined in terms of p_k, r_k and A. Show that α_k is the solution of a $n \times 1$-dimensional minimization problem of the form

$$\min_{\alpha\in\mathbb{R}} \|v_1^{(k)} - v_2^{(k)}\alpha\|^2$$

expressed using its normal equations. In your solution, identify the vector norm used to define the $n \times 1$-dimensional minimization problem, give $v_1^{(k)}$ and $v_2^{(k)}$, and show how α_k arises from the associated normal equations.

Problem 5.2

Let $A = Q\Lambda Q^T$ be a symmetric positive definite matrix where Q is an orthogonal matrix and Λ is a diagonal matrix whose diagonal elements are positive and also are the eigenvalues of A. Define

$$\tilde{x} = Q^T x \quad \text{and} \quad \tilde{b} = Q^T b$$
$$Ax = b \quad \text{and} \quad \Lambda\tilde{x} = \tilde{b}$$

Given x_0 and $\tilde{x}_0$, define the sequence x_k as the sequence of vectors produced by steepest descent applied to $Ax = b$ and the sequence $\tilde{x}_k$ as the sequence of vectors produced by steepest descent applied to $\Lambda\tilde{x} = \tilde{b}$.

Let $e^{(k)} = x_k - x$ and $\tilde{e}^{(k)} = \tilde{x}_k - \tilde{x}$. Show that if $\tilde{x}_0 = Q^T x_0$ then

$$\|e^{(k)}\|_2 = \|\tilde{e}^{(k)}\|_2, \quad k > 0$$

$$\|r_k\|_2 = \|\tilde{r}_k\|_2, \quad k > 0.$$

Also, what is the relationship between the stepsizes α_k and $\tilde{\alpha}_k$ for the x_k and $\tilde{x}_k$ iterations respectively.

Problem 5.3

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix with eigendecomposition $A = Q\Lambda Q^T$ where Q is an orthogonal matrix and Λ is a diagonal matrix whose diagonal elements are positive and also are the eigenvalues of A. Consider solving the linear system $Ax = b$ with solution $x_* = A^{-1}b$. using the general descent method.

(5.3.a) Show that for the choice of stepsize α_k used in the method we have $r_{k+1}^T p_k = 0$, i.e., $r_{k+1} \perp p_k$ in the Euclidean inner product, where $r_{k+1} = b - Ax_{k+1}$ is the residual vector for x_{k+1}.

(5.3.b) Suppose we take a direction vector p_k such that $p_k \perp r_k$, where $r_k = b - Ax_k$ is the residual vector for x_k. How does this affect the iteration?

(5.3.c) A matrix polynomial of degree $k+1$ can be defined as $P_{k+1}(A) = \nu_0 I + \nu_1 A + \cdots + \nu_k A^k + \nu_{k+1} A^{k+1}$ where the ν_i are real scalars. When analyzing iterative methods for linear systems the matrix polynomial can often be expressed in the more specific product form of degree $k+1$

$$P_{k+1}(A) = \prod_{i=0}^{k} (I - \gamma_i A) \tag{1}$$

where the γ_i are real scalars. Consider solving $Ax = b$ using the Steepest Descent method, i.e., the general descent method with $p_k = r_k$. Show that the residual at step $k+1$, $r_{k+1} = b - Ax_{k+1}$ can be written as

$$r_{k+1} = P_{k+1}(A) r_0$$

where $r_0 = b - Ax_0$ and $P_{k+1}(A)$ has the product form of (1). Be specific about relating the γ_i to parameters in the Steepest Descent sequence.

(5.3.d) Assuming $\tilde{x}_k = Q^T x_k$, $k \geq 0$ and $\tilde{b} = Q^T b$, what matrix polynomial relates $\tilde{r}_{k+1} = \tilde{b} - \Lambda \tilde{x}_{k+1}$ and $\tilde{r}_0$ for the Steepest Descent method?

Problem 5.4

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite tridiagonal matrix, i.e., its elements are 0 when not on the main diagonal or first superdiagonal or first subdiagonal. For $n = 6$, A would have the form

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 & 0 & 0 & 0 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & 0 & 0 & 0 \\ 0 & \alpha_{32} & \alpha_{33} & \alpha_{34} & 0 & 0 \\ 0 & 0 & \alpha_{43} & \alpha_{44} & \alpha_{45} & 0 \\ 0 & 0 & 0 & \alpha_{54} & \alpha_{55} & \alpha_{56} \\ 0 & 0 & 0 & 0 & \alpha_{65} & \alpha_{66} \end{pmatrix}$$

where $\alpha_{ij} = \alpha_{ji}$. Consider solving the linear system $Ax = b$ with solution $x_* = A^{-1}b$. using the general descent method.

Determine the computational complexity, i.e., what are the number of storage locations and the number of computations, for the method. Be sure to give the numbers for each major computation done in each iteration and for the matrix and any vectors required. Express the totals as

$$Cn^d + O(n^{d-1}) \text{ computations and } \tilde{C}n^{\tilde{d}} + O(n^{\tilde{d}-1}) \text{ locations.}$$

Problem 5.5

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix with eigendecomposition $A = Q\Lambda Q^T$ where Q is an orthogonal matrix and Λ is a diagonal matrix whose diagonal elements are positive and also are the eigenvalues of A. Consider solving the linear system $Ax = b$ with solution $x_* = A^{-1}b$. using the Steepest Descent method.

(5.5.a) Suppose the n eigenvalues of A all have the same value, i.e., $\lambda_{1,1} = \lambda_{2,2} = \ldots = \lambda_{n,n} = \mu > 0$. What behavior does this cause for the iteration from the Steepest Descent method for all $x_0 \in \mathbb{R}^n$?

(5.5.b) Now suppose the n eigenvalues of A on take on two distinct values, i.e., $\lambda_{1,1} = \lambda_{2,2} = \ldots = \lambda_{s,s} = \mu_1 > 0$ and $\lambda_{s+1,s+1} = \lambda_{s+2,s+2} = \ldots = \lambda_{n,n} = \mu_2 > 0$ with $\mu_1 \neq \mu_2$. Does the behavior you identified when all eigenvalues had the same value still occur? Justify your answer.

(5.5.c) For the situation where $\mu_1 \neq \mu_2$ are the only values taken on by the $\lambda_{i,i}$, relate the stepsize α_k used to compute $x_{k+1} = x_k + r_k\alpha_k$ in the Steepest Descent method to the μ_j and the residual vector r_k.

Problem 5.6

5.6.a

If CG is used to solve $Ax = b$ where A is symmetric positive definite then the iterates and errors have the form

$$x_k = x_0 + \alpha_0 d_0 + \alpha_1 d_1 + \ldots + \alpha_{k-1} d_{k-1} = x_{k-1} + \alpha_{k-1} d_{k-1}$$

$$e^{(k)} = x^* - x_k, \quad x^* = A^{-1} b$$

$$e^{(0)} = \alpha_0 d_0 + \alpha_1 d_1 + \ldots + \alpha_{n-1} d_{n-1}, \quad \alpha_i = \frac{\langle e^{(0)}, d_i \rangle_A}{\langle d_i, d_i \rangle_A}$$

$$\langle d_i, d_j \rangle_A = d_i^T A d_j = 0 \text{ for } i \neq j, \quad \langle d_i, d_i \rangle_A = d_i^T A d_i = \|d_i\|_A^2 \neq 0$$

i.e., the vectors $\{d_0, \ldots, d_{n-1}\}$ are A-orthogonal.

It can be shown that taking an arbitrary x_0 and $d_0 = r_0 = b - Ax_0$ that we have the spaces $\mathcal{S}_k$ for $k = 0, \ldots, n-1$ with multiple bases and satisfying the conditions

$$\mathcal{S}_k = span\,[d_0, d_1, \ldots, d_{k-1}, d_k] = span\,[d_0, d_1, \ldots, d_{k-1}, r_k] = span\,[r_0, r_1, \ldots, r_{k-1}, r_k]$$

$$r_k^T d_j = 0, \quad j = 0, \ldots, k-1$$
$$r_i^T r_j = 0, \quad i \neq j, \quad 0 leq i, j \leq n-1$$

$$x_{k+1} = x_0 + z_k = x_k + \alpha_k d_k, \quad z_k \in \mathcal{S}_k.$$

It is straightforward to show that for CG we have

$$r_1^T d_0 = r_1^T r_0 = 0$$

$$span\,[d_0, d_1] = span\,[r_0, r_1] = span\,[r_0, Ar_0]\,.$$

Use the definitions and properties of CG given above and assume the induction hypothesis,

$$\mathcal{S}_{k-1} = span\left[r_0, Ar_0, \ldots, A^{k-2} r_0, A^{k-1} r_0\right]$$

to show that

$$\mathcal{S}_k = span\left[r_0, Ar_0, \ldots, A^{k-1} r_0, A^k r_0\right]\,.$$

Hint: Consider the recurrence used in the efficient CG implementation to update r_{k-1} to r_k which relates r_{k-1}, r_k, d_{k-1} and A.

5.6.b

Show that x_k generated by CG satisfies

$$\|e^{(k)}\|_A^2 \leq \min_{x \in x_0 + \mathcal{S}_{k-1}} \|x^* - x\|_A^2.$$

(In fact, for CG it is a strict inequality but you need not prove that.)

Problem 5.7

Suppose A is symmetric positive definite matrix and the system $Ax = b$ with solution $x^* = A^{-1}b$ is to be solved by Steepest Descent and CG. An approximation of x^* , denoted v, is said to be accurate to d decimal digits if

$$\frac{\|x^* - v\|_A}{\|x^*\|_A} \leq 10^{-d}$$

where accuracy is measured using the A-norm in this case.

5.7.a. Suppose A is symmetric positive definite with a condition number of 10. Determine an expression for a lower bound on the number of iterations of Steepest Descent would be required to guarantee 6 places of accuracy in the solution of $Ax = b$ assuming that x_0 was accurate to 3 decimal digits?

5.7.b. Suppose all you know about A is its condition number. Would you expect Conjugate Gradient to be guaranteed to achieve the same accuracy as Steepest Descent in fewer steps than the the number you determined for the previous part of the question? If so what is the relationship between the two number of steps? If not, why not?

5.7.c. What other information about A would you want to know to show that the number of steps required by Conjugate Gradient to guarantee a given accuracy is less than the number of steps based on only the condition number?

Problem 5.8

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix, $C \in \mathbb{R}^{n\times n}$ be a symmetric nonsingular matrix, and $b \in \mathbb{R}^n$ be a vector. The matrix $M = C^2$ is therefore symmetric positive definite. Also, let $\tilde{A} = C^{-1}AC^{-1}$ and $\tilde{b} = C^{-1}b$.

The preconditioned Steepest Descent algorithm to solve $Ax = b$ is:

A, M are symmetric positive definite
x_0 arbitrary; $r_0 = b - Ax_0$; solve $Mz_0 = r_0$

do $k = 0, 1, \ldots$ until convergence

$$
\begin{aligned}
&w_k = Az_k \\
&\alpha_k = \tfrac{z_k^T r_k}{z_k^T w_k} \\
&x_{k+1} \leftarrow x_k + z_k \alpha_k \\
&r_{k+1} \leftarrow r_k - w_k \alpha_k \\
&\text{solve } M z_{k+1} = r_{k+1}
\end{aligned}
$$

end

The Steepest Descent algorithm to solve $\tilde{A}\tilde{x} = \tilde{b}$ is:

$\tilde{A}$ is symmetric positive definite
$\tilde{x}_0$ arbitrary; $\tilde{r}_0 = \tilde{b} - \tilde{A}\tilde{x}_0$; $\tilde{v}_0 = \tilde{A}\tilde{r}_0$

do $k = 0, 1, \ldots$ until convergence

$$
\begin{aligned}
&\tilde{\alpha}_k = \tfrac{\tilde{r}_k^T \tilde{r}_k}{\tilde{r}_k^T \tilde{v}_k} \\
&\tilde{x}_{k+1} \leftarrow \tilde{x}_k + \tilde{r}_k \tilde{\alpha}_k \\
&\tilde{r}_{k+1} \leftarrow \tilde{r}_k - \tilde{v}_k \tilde{\alpha}_k \\
&\tilde{v}_{k+1} \leftarrow \tilde{A}\tilde{r}_{k+1}
\end{aligned}
$$

end

Show that given the appropriate consistency between initial guesses the preconditioned steepest descent recurrences to solve $Ax = b$ can be derived from the steepest descent recurrences to solve $\tilde{A}\tilde{x} = \tilde{b}$.