Study Problems 5 Foundations of Computational Math 1 Fall 2024

These study questions concern some of the basic properities of Steepest Descent, Conjugate Gradient and members of the Richardson's family of stationary methods. They build on earlier results in earlier study questions, class notes and statements made in the lectures.

Problem 5.1

Let $A \in \mathbb{R}^{n\times n}$ be symmetric postive definite with an eigendecompositon $A = Q\Lambda Q^T$ with $Q \in \mathbb{R}^{n\times n}$ and orthogonal matrix, i.e., $Q^TQ = QQ^T = I$, and $\Lambda \in \mathbb{R}^{n\times n}$ a diagonal matrix with positive diagonal elements $\lambda_i = e_i^T \Lambda e_i > 0$.

Consider the two systems $Ax = b$ and $\Lambda\tilde{x} = \tilde{b}$ with $Q\tilde{x} = x$ and $Q\tilde{b} = b$. The iterations defined by applying Steepest Descent (SD) to each are

$$x_{k+1} = x_k + \alpha_k r_k, \quad r_k = b - Ax_k, \quad \alpha_k = \frac{r_k^T r_k}{r_k^T A r_k}$$

$$\tilde{x}_{k+1} = \tilde{x}_k + \tilde{\alpha}_k \tilde{r}_k, \quad \tilde{r}_k = \tilde{b} - \Lambda\tilde{x}_k, \quad \tilde{\alpha}_k = \frac{\tilde{r}_k^T \tilde{r}_k}{\tilde{r}_k^T \Lambda \tilde{r}_k}$$

given x_0 and $Q\tilde{x}_0 = x_0$. The elements of the vectors with the tildes are the coefficients of the corresponding vectors without the tildes with respect to the basis of eigenvectors given by the columns of Q.

We have shown in other problems that the two iterations are essentially equivalent in the behavior of the norms of the error and residual at each step. It is also known that $\alpha_k = \tilde{\alpha}_k$ and that α_k^{-1} can be written as a weighted average of the eigenvalues of A with weights determined by r_k.

(5.1.a) Consider applying SD to $Ax = b$. Derive a sufficient condition on A so that for any x_0 convergence to $A^{-1}b$ occurs in one step, i.e.,

$$A^{-1}b = x_1 = x_0 + \alpha_0 r_0.$$

(5.1.b) Is the condition also a necessary condition for convergence of SD in one step for any x_0?

(5.1.c) Does the condition imply that the stationary Richardson's method without preconditioning, $x_{k+1} = x_k + \alpha r_k$, converges in one step?

(5.1.d) Does the condition imply that CG without preconditioning $x_{k+1} = x_k + \alpha r_k$ converges in one step?

Problem 5.2

Suppose we are to solve $Ax = b$ where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite using a method based like Steepest Descent or CG that is based on reducing the error

$$E(x) = \|x - x_*\|_A^2$$

where $x_* = A^{-1}$. Recall, that it is known that x_* is also the unique minimizer of

$$f(x) = \frac{1}{2}x^T Ax - b^T x.$$

Each step of the standard methods chooses a direction p_k and then optimizes the choice of stepsize α_k so that $x_{k+1} = x_k + \alpha_k p_k$ is a minimum of $f(x_k + \alpha p_k)$ with respect to α, i.e., it minimizes f along a line defined by p_k.

5.2.a. Suppose that the particular method is of the form $x_{k+1} = x_k + \alpha_k p_k$ where α_k is chosen so x_{k+1} is a minimum of $f(x_k + \alpha p_k)$ with respect to α, i.e., it minimizes f along a line defined starting at x_k and moving in the direction of p_k. Derive an expression for $f(x_k + \alpha p_k)$ of the form

$$\phi_k(\alpha) = f(x_k + \alpha p_k) = f(x_k) + \omega_k(\alpha)$$

where $\omega_k(\alpha)$ a scalar polynomial in α with the coefficients defined in terms of p_k, r_k, x_k, and A.

5.2.b. What condition on p_k is required such that $\alpha > 0$ can be chosen so that $f(x_k + \alpha p_k) < f(x_k)$?

5.2.c. Derive the expression for α_k for any given p_k in the iteration that minimizes $f(x_k + \alpha p_k)$ for $\alpha > 0$. Is this formula consistent with what is used for Steepest Descent, i.e., when $p_k = r_k$?

5.2.d. Show that for this choice of α_k we have $r_{k+1}^T p_k = 0$, i.e., $r_{k+1} \perp p_k$.

5.2.e. Consider the optimal value α_k for a given p_k. For what range of α is $f(x_k + \alpha p_k) < f(x_k)$, i.e., consider $\alpha = \sigma\alpha_k$ for $0 \le \sigma \le \sigma_{max}$ and determine σ_{max}.

Problem 5.3

Recall, we have the two convergence theorems for A symmetric positive definition for a stationary iteration $x_{k+1} = x_k + P^{-1} r_k$ that depend on whether P is also symmetric positive definite.

1. In general, if $M = P + P^T - A$ is positive definite then the iteration converges.
2. A corollary says that if P is also symmetric positive definite then if $M = 2P - A$ is positive definite the iteration converges.

The corollary can be proven directly without appealing to the first theorem. This problem considers that proof.

5.3.a

The following basic lemma that relates a spectral radius to the definiteness of M is a key to proving this result.

Lemma. *Let $B \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. The matrix $M = 2I - B$ is positive definite if and only if $\rho(I - B) < 1$.*

Prove the lemma.

5.3.b

Prove the following theorem.

Theorem 1. *Let $A \in \mathbb{R}^{n\times n}$ and $P \in \mathbb{R}^{n\times n}$ be symmetric positive definite matrices. Also assume that $P = CC^T$ where $C \in \mathbb{R}^{n\times n}$ is nonsingular.*

$M = 2P - A$ is positive definite if and only if $\rho(I - P^{-1}A) < 1$.

Problem 5.4

5.4.a

Let $A = D - L - U \in \mathbb{R}^{n\times n}$ be a nonsingular matrix, where $-L$ is the matrix of strictly lower triangular elements and $-U$ is the matrix of strictly upper triangular elements. Recall the three methods and their preconditioners

- Gauss-Seidel (forward): $P_{gs} = D - L$
$$x_{k+1} = x_k + P_{gs}^{-1} r_k = x_k + (D - L)^{-1} r_k$$
- Gauss-Seidel (backward): $P_{bgs} = D - U$
$$x_{k+1} = x_k + P_{bgs}^{-1} r_k = x_k + (D - U)^{-1} r_k$$
- Symmetric Gauss-Seidel: $P_{sgs} = (D - L)D^{-1}(D - U)$
$$x_{k+1} = x_k + P_{sgs}^{-1} r_k = x_k + (D - U)^{-1} D (D - L)^{-1} r_k.$$

5.4.a. Show that one iteration of Symmetric Gauss-Seidel is equivalent to one iteration of forward Gauss-Seidel followed by one iteration of backward Gauss-Seidel.

5.4.b. Now assume $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and consider solving the linear system $Ax = b$. Show that the Forward Gauss-Seidel iteration converges for any x_0.

5.4.c. Again assume $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix and consider solving the linear system $Ax = b$. Show that the Symmetric Gauss-Seidel iteration converges for any x_0.

Problem 5.5

(5.5.a) Suppose you are to solve $Ax = b$ where A is known to be nonsingular via an iterative method. Which, if any, of the iterative methods, Jacobi, (forward) Gauss-Seidel, Symmetric Gauss-Seidel, Steepest Descent and CG, would converge if

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}?$$

(5.5.b) In order to have a unique solution for $Ax = b$ the matrix A must be nonsingular. If Gauss-Seidel is to converge we must have the spectral radius $\rho(G_{gs}) < 1$ where G_{gs} is the iteration matrix defining Gauss-Seidel. Must G_{gs} be nonsingular? If so, explain why. If not, i.e., if G_{gs} can be singular, identify a vector in its null space.

Problem 5.6

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 $P = C^2$ 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, P are symmetric positive definite
> x_0 arbitrary; $r_0 = b - Ax_0$; solve $Pz_0 = r_0$
>
> do $k = 0, 1, \ldots$ until convergence
>
> $$w_k = Az_k$$
> $$\alpha_k = \frac{z_k^T r_k}{z_k^T w_k}$$
> $$x_{k+1} \leftarrow x_k + z_k\alpha_k$$
> $$r_{k+1} \leftarrow r_k - w_k\alpha_k$$
> solve $Pz_{k+1} = r_{k+1}$
>
> 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

$\tilde{\alpha}_k = \frac{\tilde{r}_k^T \tilde{r}_k}{\tilde{r}_k^T \tilde{v}_k}$
$\tilde{x}_{k+1} \leftarrow \tilde{x}_k + \tilde{r}_k \tilde{\alpha}_k$
$\tilde{r}_{k+1} \leftarrow \tilde{r}_k - \tilde{v}_k \tilde{\alpha}_k$
$\tilde{v}_{k+1} \leftarrow \tilde{A} \tilde{r}_{k+1}$

end

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 $\tilde{A}\tilde{x} = \tilde{b}$.