Study Problems 3 Numerical Linear Algebra 1 Fall 2024

These are study questions and you need not submit solutions. Solutions will be posted on the class webpage. The questions span a range of complexity and topics and therefore contain some that are here only for providing more details on classic results mentioned in the notes and that you will encounter in the literature. For those, you are not expected to duplicate that effort on any graded activity for this class.

Problem 3.1

Let $T \in \mathbb{R}^{n\times n}$ be a symmetric tridiagonal matrix, i.e., $e_i^T T e_j = e_j^T T e_i$ and $e_i^T T e_j = 0$ if $j < i-1$ or $j > i+1$. Consider $T = QR$ where $R \in \mathbb{R}^{n\times n}$ is an upper triangular matrix and $Q \in \mathbb{R}^{n\times n}$ is an orthogonal matrix.

Recall, the nonzero structure of R was derived in class and shown to be $e_i^T R e_j = 0$ if $j < i$ (upper triangular assumption) or if $j > i+2$, i.e, nonzeros are restricted to the main diagonal and the first two superdiagonals.

3.1.a Show that Q has nonzero structure such that $e_i^T Q e_j = 0$ if $j < i-1$, i.e., Q is upper Hessenberg.

3.1.b Show that $T_+ = RQ$ is a symmetric triagonal matrix.

3.1.c Prove the Lemma in the class notes that states that choosing the shift $\mu = \lambda$, where λ is an eigenvalue of T, results in a reduced T_+ with known eigenvector and eigenvalue.

Problem 3.2

(Golub and Van Loan Problem 8.3.1. (3rd Ed.) p. 423, Golub and Van Loan Problem 8.3.1. (4th Ed.) p. 465)

Suppose λ is an eigenvalue of a symmetric tridiagonal matrix T. Show that if λ has algebraic mulitplicity k, then at least $k-1$ of T's subdiagonal elements are zero.

Problem 3.3

Suppose $\|v\|_\nu$ is a given vector norm on $\mathbb{R}^n$. The vector norm induces a matrix norm $\|A\|_\alpha$ on $\mathbb{R}^{n\times n}$ by a maximization-based definition.

3.3.a Show that the vector and induced matrix norms satisfy for any $A \in \mathbb{R}^{n\times n}$:

$$\forall v \in \mathbb{R}^n, \quad \|Av\|_\nu \leq \|A\|_\alpha \|v\|_\nu$$

3.3.b Show that for any pair of matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ the matrix norm satisfies

$$\|AB\|_\alpha \leq \|A\|_\alpha \|B\|_\alpha$$

Problem 3.4

Suppose $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix.

3.4.a Show that A is nonsingular.

3.4.b Show that all of the eigenvalues of A are real and positive and that they have corresponding real eigenvectors.

3.4.c Show that for any $v_1 \in \mathbb{R}^n$ and $v_2 \in \mathbb{R}^n$, the function $\langle v_1, v_2\rangle_A = v_2^T A v_1$ is an inner product.

3.4.d It is known that if A is symmetric positive definite then it has a symmetric positive definite square root, $A^{1/2} \in \mathbb{R}^{n\times n}$, such that $A = A^{1/2}A^{1/2} = A^{1/2}A^{T/2} = \left(A^{1/2}\right)^2$. It is also known that an inner product induces a vector norm by $\|v\|_A^2 = \langle v, v\rangle_A = v^T A v$. Prove that $\|v\|_A$ is a vector norm using an alternate approach that relates $\|v\|_A$ to $\|\tilde{v}\|_2$ where $\tilde{v}$ is unique for each $v \in \mathbb{R}^n$ and then exploits that it is known $\|\tilde{v}\|_2$ is a vector norm, i.e. it satisfies the required properties.

Problem 3.5

Suppose $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. Recall, that A has several well known factorizations due to symmetry and positive definiteness.

1. $A = A^T$ implies that $A = Q\Lambda Q^T$ where $Q^TQ = QQ^T = I \in \mathbb{R}^{n\times n}$ and $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{R}^{n\times n}$, i.e., Λ is a real diagonal matrix. (Schur Decomposition)

2. If A is also positive definite then $A = LDL^T$ where $L \in \mathbb{R}^{n\times n}$ is a unit lower triangular matrix and $D = diag(\delta_1, \delta_2, \ldots, \delta_n) \in \mathbb{R}^{n\times n}$ with $\delta_i > 0$.

3. If A is also positive definite then $A = \tilde{L}\tilde{L}^T$ where $L \in \mathbb{R}^{n\times n}$ is a nonsingular lower triangular matrix with $e_i^T \tilde{L} e_i > 0$. (Cholesky factorization).

3.5.a Show that if A is symmetric then $\|A\|_2 = \rho(A)$ where $\rho(A)$ is the spectral radius of A. (Note there is no assumption of positive definite in this item.)

3.5.b Show if A is symmetric then that $\|A^{-1}\|_2 = 1/|\lambda_{min}|$ where λ_{min} is the eigenvalue of A that has the minimal magnitude. (Note there is no assumption of positive definite in this item.)

3.5.c Show that if $C \in \mathbb{R}^{n \times n}$ is nonsingular then CC^T is a symmetric positive definite matrix.

3.5.d Show that if A is symmetric positive definite then it has a symmetric positive definite square root, $A^{1/2} \in \mathbb{R}^{n \times n}$, such that $A = A^{1/2}A^{1/2} = A^{1/2}A^{T/2} = \left(A^{1/2}\right)^2$.

3.5.e Show that if A is symmetric positive definite with maximum and minimum eigenvalues λ_{max} and λ_{min} then

$$\forall w \in \mathbb{R}^n, \;\; 0 < \lambda_{min} \leq \frac{w^T A w}{w^T w} \leq \lambda_{max}$$

Problem 3.6

Consider solving a linear system $Ax = b$ where A is symmetric positive definite using steepest descent.

3.6.a

Suppose you use steepest descent without preconditioning. Show that the residuals, r_k and r_{k+1} are orthogonal for all k.

3.6.b

Suppose you use steepest descent with preconditioning. Are the residuals, r_k and r_{k+1} orthogonal for all k? If not is there any vector from step k that is guaranteed to be orthogonal to r_{k+1}?

Problem 3.7

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 3.8

Let $A \in \mathbb{R}^{n \times k}$, $x \in \mathbb{R}^k$, and $b \in \mathbb{R}^n$ with the columns of A linearly independent and consider the linear least squares problem

$$\min_{x \in \mathbb{R}^k} \|b - Ax\|_2$$

3.8.a. Show that $N = A^T A \in \mathbb{R}^{k \times k}$ is a symmetric positive definite matrix

3.8.b. Suppose that n and k are both very large and that A is very sparse, i.e., a small number of nonzero elements much less than k is in each row. Show how you would use CG without preconditioning to solve for the solution of the least squares problem x_{min} in a computationally efficient manner. Comment on the complexity of one step of your algorithm in terms of order of operations (you need not worry about the multiplicative constant in the order expressions).

Problem 3.9

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 $\mathbb{R}$

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

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

3.9.b What are the gradients $\nabla E(x)$ and $\nabla f(x)$?

Problem 3.10

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.

3.10.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.,

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

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

3.10.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?

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

Problem 3.11

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.

3.11.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.

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

3.11.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 $p_k = r_k$?

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

3.11.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 $0 \leq \sigma \leq \sigma_{max}$ and determine σ_{max}.

Problem 3.12

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 $P z_{k+1} = r_{k+1}$

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