Review Sheet for Final Exam

Read also Review Sheet for Midterm Exam

Terminology

Can you explain these words?

1. permutation
2. transposition
3. sign of a permutation
4. determinant
5. minors and cofactor
6. rotation matrices
7. classical adjoint
8. eigenvalue and eigenvector
9. trace
10. characteristic equation and characteristic polynomial
11. algebraic multiplicity
12. eigenspace and geometric multiplicity
13. eigenbasis
14. diagonalizable
15. complex conjugate
16. polar form, modulus, and argument
17. inner product for complex vectors

18. orthogonally diagonalizable
19. adjoint, selfadjoint, and unitary
20. quadratic form
21. positive definite, positive semidefinite, and indefinite
22. principal submatrices
23. principal axes
24. singular values
25. matrix norm
26. condition number
27. Tikhonov regularization
28. truncated SVD

Permutation: A permutation is a mapping σ of n objects, say the numbers $1, 2, \ldots, n$ onto themselves. Like all functions, permutations can be composed. Being onto, they are one-to-one and so can be inverted. For example, consider $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. This means $\sigma(1) = 3$, $\sigma(2) = 1$, and $\sigma(3) = 2$. We can also write the same permutation as $\sigma = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$. In general, we have $\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}$, where $i_1, \ldots, i_n$ are taken from a sequence of $1, \ldots, n$ without repetition. This means $\sigma(1) = i_1, \sigma(2) = i_2, \ldots, \sigma(n) = i_n$.

Transposition: A transposition is a permutation that exchanges two numbers. For example, $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 3 & 2 \\ 3 & 1 & 2 \end{pmatrix}$ are transpositions. Any permutation can be expressed as a composition of transpositions although the decomposition is not unique. For example, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$.

Sign of a permutation: When we express a permutation σ as a composition of transpositions, the parity (even or odd) of the number of transpositions is unique. We define the sign or signature of σ as $\operatorname{sgn}(\sigma) = 1$ for even σ and $\operatorname{sgn}(\sigma) = -1$ for odd σ. For example, $\operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = 1$ and $\operatorname{sgn} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = -1$.

Determinant: For an $n \times n$ matrix A, the determinant is defined as $\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \sum_\sigma \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)}$. For example, consider the case that $n = 2$ and $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. We have two permutations: $\sigma_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ and $\sigma_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. The determinant is calculated as $\det A = \operatorname{sgn}(\sigma_1) a_{1\sigma_1(1)} a_{2\sigma_1(2)} + \operatorname{sgn}(\sigma_2) a_{1\sigma_2(1)} a_{2\sigma_2(2)} = (+1) a_{11} a_{22} + (-1) a_{12} a_{21}$.

Minor and cofactor: For an $n \times n$ matrix A, we obtain an $(n-1) \times (n-1)$ matrix A_{ij} by omitting the ith row and jth column of A. The determinant $\det(A_{ij})$ is called a minor of A. Then $(-1)^{i+j} \det(A_{ij})$ is called a cofactor of A.

Rotation matrices: An orthogonal $n \times n$ matrix A with $\det A = 1$ is called a rotation matrix, and the linear transformation $T(\vec{x}) = A\vec{x}$ is called a rotation.

Classical adjoint: The classical adjoint $\operatorname{adj}(A)$ is the $n \times n$ matrix whose ijth entry is a cofactor $(-1)^{i+j} \det(A_{ji})$.

Eigenvalue and eigenvector: Consider an $n \times n$ matrix A. A nonzero vector $\vec{v} \in \mathbb{R}^n$ is called an eigenvector of A if $A\vec{v} = \lambda\vec{v}$, where λ is a scalar. This λ is called the eigenvalue associated with the eigenvector $\vec{v}$.

Trace: The sum of the diagonal entries of a square matrix A is called the trace of A, denoted by $\operatorname{tr} A$.

Characteristic equation and characteristic polynomial: For an $n \times n$ matrix A, the polynomial $f_A(\lambda) = \det(A - \lambda I_n)$ is called the characteristic polynomial of A. This polynomial of degree n is written as $f_A(\lambda) = (-\lambda)^n + (\operatorname{tr} A)(-\lambda)^{n-1} + \cdots + \det A$. The equation $f_A(\lambda) = 0$ is called the characteristic equation.

Algebraic multiplicity: An eigenvalue λ_0 of A is said to have algebraic multiplicity k if λ_0 is a root of multiplicity k of $f_A(\lambda)$, meaning that we can write $f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$ for a polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$.

Eigenspace and geometric multiplicity: Consider an eigenvalue λ of an $n \times n$ matrix A. Then $E_\lambda = \ker(A - \lambda I_n)$ is called the eigenspace associated with λ. The geometric multiplicity of λ is $\dim(E_\lambda)$. By the rank-nullity theorem, we have $\dim(E_\lambda) = n - \operatorname{rank}(A - \lambda I_n)$.

Eigenbasis: A basis of $\mathbb{R}^n$ consisting of eigenvectors of an $n \times n$ matrix A is called an eigenbasis for A.

Diagonalizable: An $n \times n$ matrix A is called diagonalizable if A is similar to a diagonal matrix D, i.e., if there exists an invertible $n \times n$ matrix S such that $S^{-1}AS$ is diagonal.

Complex conjugate The (complex) conjugate of a complex number $z = a + ib$ is defined by $\bar{z} = a - ib$.

Polar form, modulus, and argument: The representation $z = re^{i\theta} = z(\cos\theta + i\sin\theta)$ is called the polar form of the complex number z. The length r is called the modulus of z, denoted by $|z|$, and the angle θ is called an argument of z.

Inner product for complex vectors: For vectors $\vec{x}, \vec{y} \in \mathbb{C}^n$, the inner product is defined by $\langle \vec{x}, \vec{y} \rangle = x_1\bar{y_1} + x_2\bar{y_2} + \cdots + x_n\bar{y_n}$.

Orthogonally diagonalizable: A matrix A is said to be orthogonally diagonalizable if there exists an orthogonal S such that $S^{-1}AS = S^TAS$ is diagonal.

Adjoint, selfadjoint, and unitary: For an $n \times m$ complex matrix A, the $m \times n$ matrix $A^\dagger = \bar{A}^T$ is called the adjoint of A. Sometimes A^* is used instead of $A^\dagger$. An $n \times n$ complex matrix A is called selfadjoint if $A^\dagger = A$. An $n \times n$ complex matrix U is called unitary if $U^\dagger U = I_n$.

Quadratic form: A function $q(x_1, x_2, \ldots, x_n)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is called a quadratic form if it is a linear combination of functions of the form $x_i x_j$ (i, j may be equal). A quadratic form can be written as $q(\vec{x}) = \vec{x} \cdot A\vec{x} = \vec{x}^T A\vec{x}$, for a unique symmetric $n \times n$ matrix A, called the matrix of q.

Positive definite, positive semidefinite, and indefinite: Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix. We say that A is positive definite if $q(\vec{x})$ is positive for all nonzero $\vec{x} \in \mathbb{R}^n$, and we call A positive semidefinite if $q(\vec{x}) \geq 0$, for all $\vec{x} \in \mathbb{R}^n$. Negative definite and negative semidefinite symmetric matrices are defined analogously. We call A indefinite if q takes positive and negative values.

Principal submatrices: For a symmetric $n \times n$ matrix A, let $A^{(m)}$ be the $m \times m$ matrix obtained by omitting all rows and columns of A past the mth ($m = 1, 2, \ldots, n$). These matrices $A^{(m)}$ are called the principal submatrices of A.

Principal axes: Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix with n distinct eigenvalues. Then the eigenspaces of A are called the principal axes of q.

Singular values: The singular values of an $n \times m$ matrix A are the square roots of the eigenvalues of the symmetric $m \times m$ matrix $A^T A$, listed with their algebraic multiplicities. It is customary to denote the singular values by $\sigma_1, \sigma_2, \ldots, \sigma_m$ and to list them in decreasing order: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$.

Matrix norm: The matrix norm of A is defined as $||A|| = \max\limits_{\vec{v} \neq \vec{0}} ||A\vec{v}||/||\vec{v}||$.

For ℓ_2 norm or 2-norm, we have $||A|| = \sigma_1$. The proof is given as follows. We write $A = U\Sigma V^T$, where $V = [\vec{v}_1 \ldots \vec{v}_m]$ and the diagonal entries of Σ is $\sigma_1, \sigma_2, \ldots, \sigma_m$. Let us express $\vec{v} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$. Then we have $||A\vec{v}|| = ||\Sigma V^T \vec{v}|| = \sqrt{\sigma_1^2 c_1^2 + \cdots + \sigma_m^2 c_m^2} \leq \sqrt{\sigma_1^2 c_1^2 + \cdots + \sigma_1^2 c_m^2} = \sigma_1 ||\vec{v}||$.

For ℓ_∞ norm or ∞-norm, we have $||A||_\infty = \max_i \sum_j |a_{ij}|$ (the maximum absolute row sum). The proof is given as $||A\vec{v}||_\infty = \max_i |\sum_j a_{ij} v_j| \leq \max_i \sum_j |a_{ij}||v_j| \leq \max_i \sum_j |a_{ij}| \, ||\vec{v}||_\infty$.

Condition number: For an $n \times m$ matrix A, $\kappa = ||A|| \, ||A^+||$ is called the condition number of A. For $n = m$, $||A^+|| = ||A^{-1}||$. For 2-norm, the condition number is obtained as $\kappa = \sigma_1/\sigma_m$.

Tikhonov regularization: Consider the linear system $A\vec{x} = \vec{b}$, where $n \times m$ ($n > m$) matrix A is written as $A = U\Sigma V^T$. In the Tikhonov regularization with the Tikhonov regularization parameter $\alpha > 0$, the regularized solution is obtained as $\vec{x}^*_{\text{reg}} = A^+_{\text{reg}}\vec{b} = (A^T A + \alpha^2 I_m)^{-1} A^T \vec{b} = \frac{\sigma_1^2}{\sigma_1^2 + \alpha^2} \frac{1}{\sigma_1} \vec{v}_1 (\vec{u}_1 \cdot \vec{b}) + \cdots + \frac{\sigma_m^2}{\sigma_m^2 + \alpha^2} \frac{1}{\sigma_m} \vec{v}_m (\vec{u}_m \cdot \vec{b})$.

Truncated SVD: In the truncated SVD with the regularization parameter $\alpha > 0$, we consider only singular values larger than α. The regularized solution is obtained as $\vec{x}^*_{\text{reg}} = A^+_{\text{reg}}\vec{b} = \theta(\sigma_1 - \alpha)\frac{1}{\sigma_1}\vec{v}_1(\vec{u}_1 \cdot \vec{b}) + \cdots + \theta(\sigma_m - \alpha)\frac{1}{\sigma_m}\vec{v}_m(\vec{u}_m \cdot \vec{b})$, where $\theta(\cdot)$ is the step function ($\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ for $x \leq 0$).

Several Theorems

Theorem 6.2.1, Theorem 6.2.6, and Theorem 6.2.8

> Suppose A and B are square matrices. Then $\det(A^T) = \det A$ and $\det(AB) = (\det A)(\det B)$. If A is invertible, then $\det(A^{-1}) = 1/\det A$.

Theorem 6.2.3

> Elementary row operations for determinants: (i) If B is obtained from A by dividing a row of A by a scalar k, then $\det B = (1/k)\det A$. (ii) If B is obtained from A by a row swap, then $\det B = -\det A$. (iii) If B is obtained from A by adding a multiple of a row of A to another row, then $\det B = \det A$.

Theorem 7.1.5

> For an $n \times n$ matrix A, "A is invertible" $\iff \det A \neq 0 \iff \lambda \neq 0$.

Theorem 6.2.7 and Theorem 7.3.6

> Suppose A is similar to B. Then, (i) $f_A(\lambda) = f_B(\lambda)$. (ii) $\text{rank}(A) = \text{rank}(B)$ and $\text{nullity}(A) = \text{nullity}(B)$. (iii) A and B have the same eigenvalues with the same algebraic and geometric multiplicities. (iv) $\det A = \det B$ and $\text{tr}\, A = \text{tr}\, B$.

Theorem 6.2.10 (Laplace expansion)

> Let $\det A_{ij}$ be minors of an $n \times n$ matrix A whose entries are a_{ij}. Then the Laplace expansion down the jth column is $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$, and the Laplace expansion along the ith row is $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$.

Theorem 6.3.1 and Theorem 7.1.2

> If A is an orthogonal matrix, then $\det A = 1$ or -1. The possible real eignvalues are 1 and -1.

Theorem 6.3.3

> If A is an $n \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, then $|\det A| = ||\vec{v}_1||\,||\vec{v}_2^{\perp}|| \cdots ||\vec{v}_n^{\perp}||$.

Theorem 6.3.4 and Theorem 6.3.6

The m-volume of the m-parallelepiped defined by vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ in $\mathbb{R}^n$ is $\sqrt{\det(A^T A)}$, where A is the $n \times m$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. In particular, if $m = n$, the volume is given by $|\det A|$.

Theorem 6.3.8 (Cramer's rule)

Consider the linear system $A\vec{x} = \vec{b}$, where A is an invertible $n \times n$ matrix. The components x_i of $\vec{x}$ are $x_i = \det(A_{\vec{b},i}) / \det A$, where $A_{\vec{b},i}$ is the matrix obtained by replacing the ith column of A by $\vec{b}$.

Theorem 6.3.9

The inverse of an $n \times n$ matrix A is given by $A^{-1} = \operatorname{adj}(A) / \det A$, where $\operatorname{adj}(A)$ is the classical adjoint of A.

Theorem 7.2.1

A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ is a solution to the characteristic equation (or the secular equation) $\det(A - \lambda I_n) = 0$.

Theorem 7.2.8 and Theorem 7.5.5

For an $n \times n$ matrix A with (complex) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, listed with their algebraic multiplicities, $\det A = \lambda_1 \cdots \lambda_n$ and $\operatorname{tr} A = \lambda_1 + \cdots + \lambda_n$.

Theorem 7.3.4

(i) Consider an $n \times$ matrix A. If all bases of eigenspaces of A are concatenated, then the resulting eigenvectors $\vec{v}_1, \ldots, \vec{v}_s$ are linearly independent (s is the sum of the geometric multiplicities of the eigenvalues of A). (ii) There exists an eigenbasis for an $n \times n$ matrix A if and only if the geometric multiplicities of the eigenvalues add up to n ($s = n$).

Theorem 7.3.5

If an $n \times n$ matrix A has n distinct eigenvalues, then there exists an eigenbasis for A.

HW 8 13) (Ex. 7.3.54) (Cayley-Hamilton theorem)

Every $n \times n$ matrix A satisfies its own characteristic equation: $f_A(A) = 0$.

Theorem 7.4.3

(i) Matrix A is diagonalizable if and only if there exists an eigenbasis for A.
(ii) If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Theorem 7.5.1 (De Moivre's formula)

$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$. Euler's formula: $e^{i\theta} = \cos\theta + i\sin\theta$.

Theorem 7.5.2 (Fundamental theorem of algebra)

Any polynomial $p(\lambda)$ with complex coefficients splits, that is, it can be written as a product of linear factors $p(\lambda) = k(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$, for complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, and k.

Theorem 7.5.4

A complex $n \times n$ matrix has n complex eigenvalues if they are counted with their algebraic multiplicities.

Theorem 8.1.1 (Spectral theorem)

A matrix A is orthogonally diagonalizable if and only if A is symmetric. A complex matrix A is diagonalizable with a unitary matrix U if and only if A is selfadjoint.

Theorem 8.1.2

Consider a symmetric matrix A. If $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A with distinct eigenvalues λ_1 and λ_2, then $\vec{v}_1 \cdot \vec{v}_2 = 0$.

Theorem 8.1.3

A symmetric $n \times n$ matrix A has n real eigenvalues if they are counted with their algebraic multiplicities.

Theorem 8.2.2

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix. Let $\mathcal{B}$ be an orthonormal eigenbasis for A, with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then, $q(\vec{x}) = \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2$, where c_i are the coordinates of $\vec{x}$ with respect to $\mathcal{B}$.

Theorem 8.2.4

A symmetric matrix A is positive definite (positive semidefinite) if and only if all of its eigenvalues are positive (nonnegative).

Theorem 8.2.5

A symmetric matrix A is positive definite if and only if $\det(A^{(m)}) > 0$ for all principal submatrices $A^{(m)}$ $(m = 1, \ldots, n)$.

Theorem 8.3.4

If A is an $n \times m$ matrix of rank r, then the singular values $\sigma_1, \ldots, \sigma_r$ are nonzero, while $\sigma_{r+1}, \ldots, \sigma_m$ are zero.

Theorem 8.3.5 (Singular value decomposition)

Any $n \times m$ matrix A can be written as $A = U\Sigma V^T$, where U is an orthogonal $n \times n$ matrix, V is an orthogonal $m \times m$ matrix, and Σ is an $n \times m$ matrix whose first r diagonal entries are the nonzero singular values $\sigma_1, \ldots, \sigma_r$ of A, and all other entries are zero ($r = \text{rank}(A)$). The matrix A can also be written as $A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$, where $\vec{u}_i$ and $\vec{v}_i$ are the columns of U and V. If A is a complex $n \times m$ matrix, then we have $A = U\Sigma V^\dagger$, where U is a unitary $n \times n$ matrix, V is a unitary $m \times m$ matrix, and Σ is an $n \times m$ matrix whose first r diagonal entries are the nonzero singular values $\sigma_1, \ldots, \sigma_r$ of A, and all other entries are zero ($r = \text{rank}(A)$).

(Regularization)

Consider the linear system $A\vec{x} = \vec{b}$, where $n \times m$ $(n > m)$ matrix A is written as $A = U\Sigma V^T$. Here, $U = [\vec{u}_1 \cdots \vec{u}_n]$, $V = [\vec{v}_1 \cdots \vec{v}_m]$, and the diagonal entries of Σ are $\sigma_1, \ldots, \sigma_m$. In the Tikhonov regularization with the Tikhonov regularization parameter $\alpha > 0$, the regularized solution is obtained as $\vec{x}^*_{\text{reg}} = \frac{\sigma_1^2}{\sigma_1^2+\alpha^2} \frac{1}{\sigma_1} \vec{v}_1 (\vec{u}_1 \cdot \vec{b}) + \cdots + \frac{\sigma_m^2}{\sigma_m^2+\alpha^2} \frac{1}{\sigma_m} \vec{v}_m (\vec{u}_m \cdot \vec{b})$. The vector $\vec{x}^*_{\text{reg}}$ minimizes $\varepsilon(\vec{x}) = ||A\vec{x} - \vec{b}||^2 + ||\alpha\vec{x}||^2$. In the truncated SVD with the regularization parameter $\alpha > 0$, the regularized solution is obtained as $\vec{x}^*_{\text{reg}} = \theta(\sigma_1 - \alpha) \frac{1}{\sigma_1} \vec{v}_1 (\vec{u}_1 \cdot \vec{b}) + \cdots + \theta(\sigma_m - \alpha) \frac{1}{\sigma_m} \vec{v}_m (\vec{u}_m \cdot \vec{b})$, where $\theta(\cdot)$ is the step function ($\theta(x) = 1$ for $x > 0$ and $\theta(x) = 0$ for $x \leq 0$).

Some Proofs

First go back to the previous section and think how to prove the theorem. Then, read the proof of the theorem below.

Proof of 7.5.1 Let us first show Euler's formula. By Taylor series, $e^{i\theta} = 1+i\theta+(i\theta)^2/2!+(i\theta)^3/3!+\cdots = (1-\theta^2/2!+\dots)+i(\theta-\theta^3/3!+\dots)$. On the other hand, $\cos\theta = 1-\theta^2/2!+\theta^4/4!-\dots$ and $\sin\theta = \theta-\theta^3/3!+\theta^5/5!-\dots$. Therefore $e^{i\theta} = \cos\theta + i\sin\theta$. Thus we have $(\cos\theta + i\sin\theta)^n = \left(e^{i\theta}\right)^n = e^{in\theta} = \cos(n\theta) + i\sin(n\theta)$.

Proof of Theorem 8.2.5 ($\Longrightarrow$) Let $p_+(A^{(m)})$ denote the number of positive eigenvalues of $A^{(m)}$. Consider a subspace S such that $q_m(\vec{x}) = \vec{x} \cdot A^{(m)}\vec{x} > 0$ for $\vec{x} \in S$ ($\vec{x} \neq \vec{0}$). First let us show that $p_+(A^{(m)}) = \max\dim S$. By Theorem 8.2.2, we can write $\vec{x} \cdot A^{(m)}\vec{x} = \sum_{i=1}^m \lambda_i c_i^2$. We label the eigenvalues so that $\lambda_1, \dots, \lambda_p$ are positive. Define the subspace S_+ to consist of all $\vec{x}$ for which $c_{p+1} = \cdots = c_m = 0$. Then $\dim S_+ = p$, and $\vec{x} \cdot A^{(m)}\vec{x} > 0$ for $\vec{x} \in S_+$. Hence $p_+(A^{(m)}) \leq \max\dim S$. Now suppose that $\dim S > p_+(A^{(m)})$. Let us introduce P that maps $c_i \in S$ into S_+ by setting all components $c_i = 0$ for $i > p$. Note that $\dim S_+ < \dim S$. Therefore, $\ker(P) \neq \{\vec{0}\}$ by the rank-nullity theorem. By definition of P, the first p components of a nonzero $\vec{y} \in \ker(P)$ are zero. But then we have $q_m(\vec{y}) \leq 0$, which shows that q_m is not positive on S. Therefore we conclude that $p_+(A^{(m)}) = \max\dim S$.

Next we show that $p_+(A^{(m)}) - 1 \leq p_+(A^{m-1}) \leq p_+(A^{(m)})$. Consider subspaces S and S' such that $q_m(\vec{x}) > 0$ for $\vec{x} \in S$ ($\vec{x} \neq \vec{0}$) and $q_{m-1}(\vec{x}') > 0$ for $\vec{x}' \in S'$ ($\vec{x}' \neq \vec{0}$). Define S as the subspace of $\mathbb{R}^m$ consisting of vectors $\vec{x}$ whose first component is zero, and the vector $\vec{x}'$ formed by the remaining $m-1$ components belonging to S'. Then for $\vec{x} \in S$, we have $\vec{x} \cdot A^{(m)}\vec{x} = \vec{x}' \cdot A^{(m-1)}\vec{x}'$. Therefore, $q_{m-1}(\vec{x}') > 0 \Rightarrow q_m(\vec{x}) > 0$. We obtain $p_+(A^{(m)}) \geq \dim S = \dim S' = p_+(A^{(m-1)})$. Thus the right-hand side $p_+(A^{m-1}) \leq p_+(A^{(m)})$ was shown. To show the left-hand side $p_+(A^{(m)}) - 1 \leq p_+(A^{m-1})$, we proceed in reverse. We start with a subspace S of $\mathbb{R}^m$. Denote by S'' the subspace of S consisting of $\vec{x}$ whose first component is zero. Hence $\dim S'' \geq \dim S - 1$. Denote by S' the subspace of $\mathbb{R}^{m-1}$ consisting of $\vec{x}'$ obtained by removing the first component of $\vec{x} \in S''$. Since the component removed in zero, $\vec{x} \cdot A^{(m)}\vec{x} = \vec{x}' \cdot A^{(m-1)}\vec{x}'$ holds for $\vec{x} \in S''$. Therefore, $q_m(\vec{x}) > 0 \Rightarrow q_{m-1}(\vec{x}') > 0$. We conclude that $p_+(A^{m-1}) \geq \dim S' = \dim S'' \geq \dim S - 1 = p_+(A^{(m)}) - 1$. In particular, if the eigenvalues of $A^{(m)}$ are all positive, i.e., $p_+(A^{(m)}) = m$, then

$p_+(A^{(m-1)}) = m-1$. That is, if the eigenvalues of $A = A^{(n)}$ are positive, eigenvalues of $A^{(n-1)}, \ldots, A^{(1)}$ are all positive.

According to Theorem 8.2.4, the eigenvalues of a positive definite matrix is positive. Thus, $A^{(m)}$ $(m = 1, \ldots, n)$ are all positive definite. According to Theorem 7.2.8, $\det A^{(m)}$ is the product of the eigenvalues of $A^{(m)}$. Therefore, if A is positive definite, then $\det(A^{(m)}) > 0$ for $m = 1, \ldots, n$.

($\Longleftarrow$) We note that, by replacing A with $-A$ in $p_+(A^{(m)})-1 \leq p_+(A^{m-1}) \leq p_+(A^{(m)})$, we see that $p_-(A^{(m)}) - 1 \leq p_-(A^{m-1}) \leq p_-(A^{(m)})$ also holds, where $p_-(A^{(m)})$ denotes the number of negative eigenvalues of $A^{(m)}$.

If $\det A^{(m)} > 0$ for all $m = 1, \ldots, n$, then each $A^{(m)}$ has an even number of negative eigenvalues (use Theorem 7.2.8). However, using the relation $p_-(A^{(m)}) - 1 \leq p_-(A^{m-1}) \leq p_-(A^{(m)})$, we see that the numbers of negative eigenvalues of $A^{(m)}$ and $A^{(m-1)}$ differ at most by 1. Thus $A^{(m)}$ has as many negative eigenvalues as $A^{(m-1)}$. Now $A^{(1)}$ is a 1×1 matrix and has no negative eigenvalue. Therefore all $A^{(2)}, A^{(3)}, \ldots$ have no negative eigenvalue. In particular, the eigenvalues of $A^{(n)} = A$ are all positive. This means that A is positive definite (see Theorem 8.2.4).

Proof of Regularization In the Tikhonov regularization, we modify the least-squares solution as $\vec{x}^*_{\rm reg} = A^+_{\rm reg}\vec{b} = (A^TA + \alpha^2 I_m)^{-1}A^T\vec{b}$. Using the SVD $A = U\Sigma V^T$, we obtain $\vec{x}^*_{\rm reg} = V(\Sigma^T\Sigma + \alpha^2 I_m)^{-1}\Sigma^T U^T\vec{b} =$

$$[\vec{v}_1 \cdots \vec{v}_m] \begin{bmatrix} \frac{1}{\sigma_1^2+\alpha^2} & & \\ & \ddots & \\ & & \frac{1}{\sigma_m^2+\alpha^2} \end{bmatrix} \Sigma^T [\vec{u}_1 \cdots \vec{u}_n]^T \vec{b} = \sum_{i=1}^m \frac{\sigma_i^2}{\sigma_i^2+\alpha^2}\frac{1}{\sigma_i}\vec{v}_i(\vec{u}_i \cdot \vec{b}).$$

Note that $0 < \frac{\sigma^2}{\sigma^2+\alpha^2} < 1$ for $0 < \sigma < \infty$.

Let us write $\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$. Then we obtain $\varepsilon(\vec{x}) = \sum_{i=1}^m(\sigma_i^2 + \alpha^2)c_i^2 - 2\sum_{i=1}^m \sigma_i(\vec{u}_i \cdot \vec{b})c_i + ||\vec{b}||^2$. To minimize ε, we determine c_i such that $\frac{\partial\varepsilon}{\partial c_i} = 0$. We obtain $c_i = [\sigma_i/(\sigma_i^2 + \alpha^2)]\vec{u}_i \cdot \vec{b}$. Thus $\vec{x}^*_{\rm reg}$ minimizes ε.

For the truncated SVD with the regularization parameter $\alpha > 0$, we choose k $(< m)$ such that $\sigma_k > \alpha$ and $\sigma_{k+1} \leq \alpha$. Without regularization, the least-squares solution is written as $\vec{x}^* = V(\Sigma^T\Sigma)^{-1}\Sigma^T U^T\vec{b}$. In the truncated SVD, we replace $(\Sigma^T\Sigma)^{-1}$ with the $m \times m$ matrix in which only the first k diagonal entries $1/\sigma_1^2, \ldots, 1/\sigma_k^2$ are nonzero and all other entries are zero. Then the regularized solution is obtained as $\vec{x}^*_{\rm reg} = A^+_{\rm reg}\vec{b} = \sum_{i=1}^k \frac{1}{\sigma_i}\vec{v}_i(\vec{u}_i \cdot \vec{b})$. By using the step function, we have $\vec{x}^*_{\rm reg} = \theta(\sigma_1 - \alpha)\frac{1}{\sigma_1}\vec{v}_1(\vec{u}_1 \cdot \vec{b}) + \cdots + \theta(\sigma_m - \alpha)\frac{1}{\sigma_m}\vec{v}_m(\vec{u}_m \cdot \vec{b})$.

More Problems

Go over homework problems. Here are more problems if you need. Solutions can be found in the textbook.

Chapter 6

6.1.1, 6.1.3, 6.1.5, 6.1.7, 6.1.9, 6.1.31, 6.1.33, 6.1.35, 6.1.37, 6.1.39, 6.1.41, 6.2.1, 6.2.3, 6.2.5, 6.2.7, 6.2.9, 6.3.23, 6.3.33

Chapter 7

7.1.9, 7.1.11, 7.1.13, 7.1.39, 7.1.41, 7.2.1, 7.2.3, 7.2.5, 7.2.7, 7.2.9, 7.2.11, 7.2.13, 7.3.1, 7.3.3, 7.3.5, 7.3.7, 7.3.9, 7.3.11, 7.3.13, 7.3.15, 7.3.17, 7.4.1, 7.4.3, 7.4.5, 7.4.7, 7.4.9, 7.4.11, 7.4.13, 7.4.15, 7.4.17, 7.4.19, 7.4.31, 7.4.33, 7.5.1, 7.5.3, 7.5.5

Chapter 8

8.1.1, 8.1.3, 8.1.5, 8.1.7, 8.1.9, 8.1.11, 8.2.5, 8.2.7, 8.2.23, 8.2.25, 8.3.1, 8.3.7, 8.3.9, 8.3.11, 8.3.13, 8.3.15, 8.3.17, 8.3.19, 8.3.25, 8.3.27, 8.3.29, 8.3.31, 8.3.33, 8.3.35