Chapter 7

Bilinear and Quadratic Forms

7.1 Notes

10/18:

- **Bilinear form**: A function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that

$$L(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2, \mathbf{y}) = \alpha L(\mathbf{x}_1, \mathbf{y}) + \beta L(\mathbf{x}_2, \mathbf{y}) \qquad L(\mathbf{x}, \alpha\mathbf{y}_1 + \beta\mathbf{y}_2) = \alpha L(\mathbf{x}, \mathbf{y}_1) + \beta L(\mathbf{x}, \mathbf{y}_2)$$

  - $L(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, \mathbf{y})$.
- **Quadratic form**: A bilinear form $L(\mathbf{x}, \mathbf{x})$.
  - $(\mathbf{x}, \mathbf{x})$ is a polynomial of degree 2 in $\mathbf{x}_1, \dots, \mathbf{x}_n$:

$$L(\lambda\mathbf{x}, \lambda\mathbf{x}) = (\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda^2(\mathbf{x}, \mathbf{x})$$

- We have that

$$(A\mathbf{x}, \mathbf{x}) = (A\lambda\mathbf{x}, \lambda\mathbf{x}) = \lambda^2(A\mathbf{x}, \mathbf{x}) = \sum_{j,i=1}^{n} \alpha_{j,i}\mathbf{x}_i\mathbf{x}_j$$

- The general form of a quadratic form:
  - Can any quadratic form on $\mathbb{R}^n$ be written as $(A\mathbf{x}, \mathbf{x})$?

10/20:

- Bilinear forms are linear in each argument when keeping the other fixed.
- Quadratic forms $Q(\mathbf{x}) = L(\mathbf{x}, \mathbf{x})$ are quadratic polynomials in the coordinates of x.
  - In particular, $Q(\lambda\mathbf{x}) = |\lambda|^2 Q(\mathbf{x})$.
- If Q quadratic is real, then $Q(\mathbf{x}) = (A\mathbf{x}, \mathbf{x})$ where A is some square matrix.
  - If $\mathbf{e}_1, \dots, \mathbf{e}_n$ is an orthonormal basis of $\mathbb{R}^n$, then there exists a unique $A = A^*$ such that $(A)_{ij} = L(\mathbf{e}_i, \mathbf{e}_j)$.
  - Keeping $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i, \mathbf{e}_i$ foxed, we have

$$\begin{aligned}
Q(\mathbf{x}) &= L(\mathbf{x}, \mathbf{x}) \\
&= L(\sum^n \mathbf{x}_i\mathbf{e}_i, \sum^n \mathbf{x}_j\mathbf{e}_j) \\
&= \sum_{i=1}^n \mathbf{x}_i L(\mathbf{e}_i, \sum^n \mathbf{x}_j\mathbf{e}_j) \\
&= \sum_{i,j=1}^n \mathbf{x}_i\mathbf{x}_j \underbrace{L(\mathbf{e}_i, \mathbf{e}_j)}_{A_{ij}}
\end{aligned}$$

- We have that

$$\begin{aligned}(A\mathbf{x},\mathbf{x}) &= (UDU^{-1}\mathbf{x},\mathbf{x})\\ &= (DU^{-1}\mathbf{x},U^{-1}\mathbf{x})\\ &= \sum_{i=1}^{n}\lambda_i(\underbrace{U^{-1}\mathbf{x}}_{\mathbf{y}_i})_i(\underbrace{U^{-1}\mathbf{x}}_{\mathbf{y}_i})_i\end{aligned}$$

- Can we characterize the set $\{\mathbf{x} : (A\mathbf{x},\mathbf{x}) = 1\}$?
  - Note that this set is equivalent to $\{\mathbf{y} : (D\mathbf{y},\mathbf{y}) = 1\}$ by teh above. This set is a rotation of the previous one. Ellipse?
- Positive quadratic form:
  - Q is positive definite if $Q(\mathbf{x}) > \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$ and Q is positive semidefinite if $Q(\mathbf{x}) \geq \mathbf{0}$ for all $\mathbf{x} \neq \mathbf{0}$.
  - Take a self-adjoint matrix $A = A^*$. It is positive definite if $Q(\mathbf{x}) = (A\mathbf{x},\mathbf{x})$ is positive definite.
- Theorem: If $A = A^*$, then
  1. A is positive definite if and only if all eigenvalues of A are positive.
  2. A is positive semidefinite if and only if all eigenvalues of A are nonnegative.
  3. A is negative semidefinite if and only if all eigenvalues of A are nonpositive.
  4. A is negative definite if and only if all eigenvalues of A are negative.
  5. A is indefinite if and only if the eigenvalues of A have positive and negative values.
- Theorem: $A = A^*$ is positive definite iff $\det A_k > 0$ for all $k = 1,\dots,n$ where A_k is the upper left $k \times k$ submatrix.
- Minimax representation of eigenvalues of a self-adjoint A.
  - Let E be a subspace of X where $\dim X < \infty$. We define $\operatorname{codim}(E) = \dim E^\perp$.
  - Thus, $\dim E + \operatorname{codim} E = \dim X$.
  - Theorem: Let $A = A^*$, $\lambda_1 \geq \cdots \geq \lambda_n$ eigenvalues of A. Then

$$\lambda_k = \max_{\substack{\text{E subspace}\\ \dim E=k}}\ \min_{\substack{\mathbf{x}\in E\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x}) = \min_{\substack{\text{F subspace}\\ \operatorname{codim} F=k-1}}\ \max_{\substack{\mathbf{x}\in F\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x})$$

    - Proof: A diagonal equals $(\lambda_1,\dots,\lambda_n)$.
    - An orthonormal basis of X such that $\dim E = k$, $\operatorname{codim} F = k-1$, $\dim F = n-k+1$.
    - There exists an $\mathbf{x}_0 \neq \mathbf{0}$ such that $\mathbf{x}_0 \in E \cap F$.
    - Note that if $B = B^*$, then the max and min of $(B\mathbf{x},\mathbf{x})$ over the unit sphere is the maximal and minimal eigenvalue of B.
    - Thus,

$$\min_{\substack{\mathbf{x}\in E\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x}) \leq (A\mathbf{x}_0,\mathbf{x}_0) \leq \max_{\substack{\mathbf{x}\in F\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x})$$

    - This is true for any E, F subspaces. $\dim E = k$, $\operatorname{codim} F = k-1$, $E_0 = \operatorname{span}(\mathbf{e}_1,\dots,\mathbf{e}_k)$ and $F_0 = \operatorname{span}(\mathbf{e}_k,\dots,\mathbf{e}_n)$.
    - Thus,

$$\min_{\substack{E_0\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x}) = \lambda_k = \max_{\substack{F_0\\ \|\mathbf{x}\|=1}}(A\mathbf{x},\mathbf{x})$$

- Additionally,

$$\lambda_{k_1} \leq \max{}_{\dim E=k}^{E} \min_{\mathbf{x}}(A\mathbf{x},\mathbf{x}) \leq \min_{\substack{F \\ \operatorname{codim} F=k-1}} \max_{\mathbf{x}}(A\mathbf{x},\mathbf{x}) \leq \lambda_k$$

- Corollary: Let $A = A^* = (a_{jk})_{1\leq j,k\leq n}$ with eigenvalues $\lambda_1,\dots,\lambda_n$ listed in decreasing order. Let $\tilde{A} = (a_{j,k})_{1\leq j,k\leq n-1}$ with eigenvalues $\mu_1,\dots,\mu_{n-1}$ listed in decreasing order. Then $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n$.
  - Consider $(A\mathbf{x},\mathbf{x})$ on $\{\mathbf{e}_1,\dots,\mathbf{e}_n\}$, but then restrict yourself to $\mathbf{x}\in\mathbb{R}^{n-1}$ on $\{\mathbf{e}_1,\dots,\mathbf{e}_{n-1}\}$.

7.2 Chapter 7: Bilinear and Quadratic Forms

From Treil (2017).

10/25:

- **Bilinear form** (on $\mathbb{R}^n$): A function $L(\mathbf{x},\mathbf{y})$ of two arguments $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ that is linear in each argument.
  - Linearity in each argument:

$$L(\alpha\mathbf{x}_1+\beta\mathbf{x}_2,\mathbf{y}) = \alpha L(\mathbf{x}_1,\mathbf{y})+\beta L(\mathbf{x}_2,\mathbf{y}) \qquad L(\mathbf{x},\alpha\mathbf{y}_1+\beta\mathbf{y}_2) = \alpha L(\mathbf{x},\mathbf{y}_1)+\beta L(\mathbf{x},\mathbf{y}_2)$$

- If $\mathbf{x}=(x_1,\dots,x_n)^T$ and $\mathbf{y}=(y_1,\dots,y_n)^T$, then

$$\begin{aligned}L(\mathbf{x},\mathbf{y}) &= \sum_{j,k=1}^{n} a_{j,k}x_ky_j\\ &= (A\mathbf{x},\mathbf{y})\\ &= \mathbf{y}^TA\mathbf{x}\end{aligned}$$

  where

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n}\\ \vdots & & \vdots\\ a_{n,1} & \cdots & a_{n,n}\end{pmatrix}$$

  - A is uniquely determined by L.
- **Quadratic form** (on $\mathbb{R}^n$): The diagonal of a bilinear form L, i.e., a bilinear form $Q[\mathbf{x}] = L(\mathbf{x},\mathbf{x}) = (A\mathbf{x},\mathbf{x})$.
  - Alternatively: A homogeneous polynomial of degree 2, i.e., a polynomial in $x_1,\dots,x_n$ with only ax_k^2 and cx_jx_k terms.
- There are infinitely many ways to write a quadratic form as $(A\mathbf{x},\mathbf{x})$.
  - However, there is a unique representation $(A\mathbf{x},\mathbf{x})$ where A is a (real) symmetric matrix.
- **Quadratic form** (on $\mathbb{C}^n$): A function of the form $Q[\mathbf{x}] = (A\mathbf{x},\mathbf{x})$ where A is self-adjoint.
- Lemma 7.1.1: Let $(A\mathbf{x},\mathbf{x})$ be real for all $\mathbf{x}\in\mathbb{C}^n$. Then $A=A^*$.
- To classify quadratic forms, consider the set of points $\mathbf{x}\in\mathbb{R}^n$ defined by $Q[\mathbf{x}]=1$ for some quadratic form Q.
  - If the matrix of Q is diagonal, i.e., $Q[\mathbf{x}] = a_1x_1^2+\cdots+a_nx_n^2$, then the set of points can easily be visualized.
- The standard method of diagonalizing a quadratic form is change of variables.
- Orthogonal diagonalization.

  - Let $Q[\mathbf{x}] = (A\mathbf{x}, \mathbf{x})$ in $\mathbb{F}^n$.
  - Suppose $\mathbf{y} = S^{-1}\mathbf{x}$ where S is an invertible $n \times n$ matrix. Then

$$Q[\mathbf{x}] = Q[S\mathbf{y}] = (AS\mathbf{y}, S\mathbf{y}) = (S^*AS\mathbf{y}, \mathbf{y})$$

    so in the new variables $\mathbf{y}$, the quadratic form has matrix S^*AS.
  - Thus, we can let $A = UDU^*$, choose $D = U^*AU$ as our new (diagonal) matrix, and let this matrix act on the variables $\mathbf{y} = U^*\mathbf{x}$.

- Non-orthogonal diagonalization:
  - Completing the square:
    - Eliminate all x_ix_j terms by completing the square. Then substitute in a y_k for each squared term.
  - Row/column operations:
    - Augment $(A|I)$. Row reduce A to D. Then $I \to S^*$.

10/28:

- **Silvester's Law of Inertia**: For a Hermitian matrix A (i.e., for a quadratic form $Q[\mathbf{x}] = (A\mathbf{x}, \mathbf{x})$) and any of its diagonalizations $D = S^*AS$, the number of positive, negative, and zero diagonal entries of D depends only on A, but not on a particular choice of diagonalization.
- **Positive** (subspace $E \subset \mathbb{F}^n$ corresponding to A): A subspace E such that $(A\mathbf{x}, \mathbf{x}) > 0$ for all nonzero $\mathbf{x} \in E$. *Also known as* ***A*-positive**.
- **Negative** (subspace $E \subset \mathbb{F}^n$ corresponding to A): A subspace E such that $(A\mathbf{x}, \mathbf{x}) < 0$ for all nonzero $\mathbf{x} \in E$. *Also known as* ***A*-negative**.
- **Neutral** (subspace $E \subset \mathbb{F}^n$ corresponding to A): A subspace E such that $(A\mathbf{x}, \mathbf{x}) = 0$ for all nonzero $\mathbf{x} \in E$. *Also known as* ***A*-neutral**.
- Theorem 7.3.1: Let A be an $n \times n$ Hermitian matrix, and let $D = S^*AS$ be its diagonalization by an invertible matrix S. Then the number of positive (resp. negative) diagonal entries of D coincides with the maximal dimension of an A-positive (resp. A-negative) subspace.
- Lemma 7.3.2: Let $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$. Then the number of positive (resp. negative) diagonal entries of D coincides with the maximal dimension of a D-positive (resp. D-negative) subspace.
- **Positive definite** (quadratic form Q): A quadratic form Q such that $Q[\mathbf{x}] > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
- **Positive semidefinite** (quadratic form Q): A quadratic form Q such that $Q[\mathbf{x}] \geq 0$ for all $\mathbf{x}$.
- **Negative definite** (quadratic form Q): A quadratic form Q such that $Q[\mathbf{x}] < 0$ for all $\mathbf{x} \neq \mathbf{0}$.
- **Negative semidefinite** (quadratic form Q): A quadratic form Q such that $Q[\mathbf{x}] \leq 0$ for all $\mathbf{x}$.
- **Indefinite** (quadratic form Q): A quadratic form Q for which there exist $\mathbf{x}_1, \mathbf{x}_2$ such that $Q[\mathbf{x}_1] > 0$ and $Q[\mathbf{x}_2] < 0$.
- **Positive definite** (Hermitian matrix A): A matrix A for which the corresponding quadratic form $Q[\mathbf{x}] = (A\mathbf{x}, \mathbf{x})$ is positive definite.
  - Positive semidefinite, negative definite, negative semidefinite, and indefinite Hermitian matrices are defined similarly.
- Theorem 7.4.1: Let $A = A^*$. Then
  1. A is positive definite iff all eigenvalues of A are positive.
  2. A is positive semidefinite iff all eigenvalues of A are non-negative.

3. A is negative definite iff all eigenvalues of A are negative.
4. A is negative semidefinite iff all eigenvalues of A are non-positive.
5. A is indefinite iff it has both positive and negative eigenvalues.

- **Upper left submatrix** (of A): A $k \times k$ matrix A_k composed of all entries of A from row (column) 1 through k in the same arrangement.
- Theorem 7.4.2 (Silvester's Criterion of Positivity): A matrix $A = A^*$ is positive definite if and only if $\det A_k > 0$ for all $k = 1, \ldots, n$.
  - To check if a matrix A is negative definite, check that the matrix $-A$ is positive definite.
- Theorem 7.4.3 (Minimax characterization of eigenvalues): Let $A = A^*$ be an $n \times n$ matrix and let $\lambda_1 \geq \cdots \geq \lambda_n$ be its eigenvalues taken in decreasing order. Then

$$\lambda_k = \max_{E:\dim E=k} \min_{\mathbf{x}\in E:\|\mathbf{x}\|=1} (A\mathbf{x}, \mathbf{x}) = \min_{F:\operatorname{codim} F=k-1} \max_{\mathbf{x}\in F:\|\mathbf{x}\|=1} (A\mathbf{x}, \mathbf{x})$$

- Corollary 7.4.4 (Intertwining of eigenvalues): Let $A = A^* = \{a_{j,k}\}_{j,k=1}^n$ be a self-adjoint matrix and let $\tilde{A} = \{a_{j,k}\}_{j,k=1}^{n-1}$ be its submatrix of size $(n-1) \times (n-1)$. Let $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_{n-1}$ be the eigenvalues of A and $\tilde{A}$ respectively, taken in decreasing order. Then

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n$$