MATH-GA2120 Linear Algebra II

Diagonalization of Quadratic Form
Sylvester's Law of Inertia
Cayley-Hamilton Theorem

Deane Yang

Courant Institute of Mathematical Sciences
New York University

April 29, 2024

Example (Part 1)

- Let

$$Q(e_1x + e_2y) = ax^2 + 2bxy + cy^2,$$

  where $a \neq 0$

- Then, completing the square,

$$\begin{aligned} Q(e_1x + e_2y) &= ax^2 + 2bxy + cy^2 \\ &= a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2 \end{aligned}$$

Example (Part 2)

- If

$$f_1 = e_1 \text{ and } f_2 = -\alpha e_1 + e_2,$$

then

$$\begin{aligned} Q(f_1 u + f_2 v) &= Q(e_1 u + (-\alpha e_1 + e_2)v) \\ &= Q((u - \alpha v)e_1 + ve_2) \\ &= a(u - \alpha v)^2 + 2b(u - \alpha v)v + cv^2) \\ &= au^2 + 2(-\alpha a + b)uv + (a\alpha^2 - 2b\alpha + c)v^2 \end{aligned}$$

- Therefore, if we assume $a \neq 0$ and set

$$\alpha = \frac{b}{a},$$

then

$$Q(f_1 u + f_2 v) = au^2 + \left(c - \frac{b^2}{a}\right) v^2$$

Example (Part 3)

- If

$$g_1 = pf_1 \text{ and } g_2 = qf_2,$$

  then

$$\begin{aligned} Q(g_1 s + g_2 t) &= Q(pf_1 s + qf_2 t) \\ &= (ap^2)s^2 + q^2 \left(c - \frac{b^2}{a}\right) t^2 \end{aligned}$$

- It follows that if p and q are chosen appropriately, then

$$Q(g_1 s + g_2 t) = \begin{cases} s^2 + t^2 & \text{if } a > 0 \text{ and } ac - b^2 > 0 \\ s^2 & \text{if } a > 0 \text{ and } ac - b^2 = 0 \\ s^2 - t^2 & \text{if } a > 0 \text{ and } ac - b^2 < 0 \\ -s^2 + t^2 & \text{if } a < 0 \text{ and } ac - b^2 < 0 \\ -s^2 & \text{if } a < 0 \text{ and } ac - b^2 = 0 \\ -s^2 - t^2 & \text{if } a < 0 \text{ and } ac - b^2 > 0 \end{cases}$$

Signature of Quadratic Form

- The **signature** of a diagonal matrix is (a, b, c), where a is the number of positive diagonal elements, b is the number of negative diagonal elements, and c is the number of zero diagonal elements
- **Sylvester's Law of Inertia:** Any two diagonalizations of a quadratic form has the same signature

Sylvester's Law of Inertia

- Let $Q : V \to \mathbb{R}$ be a quadratic form
- Let $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ be bases of V that diagonalize Q
- I.e., for any $v = e_k a^k = f_k b^k$,

$$\begin{aligned} Q(v) &= Q(e_k a^k) \\ &= \alpha_1 (a^1)^2 + \cdots + \alpha_n (a^n)^2 \\ &= Q(f_k b^k) \\ &= \beta_1 (b^1)^2 + \cdots + \beta_n (b^n)^2 \end{aligned}$$

- We want to show that the number of positive values in $\{\alpha^1, \ldots, \alpha^n\}$ equals the number of positive values in $\{\beta^1, \ldots, \beta^n\}$
- The same argument will also imply that the number of negative values in $\{\alpha^1, \ldots, \alpha^n\}$ equals the number of negative values in $\{\beta^1, \ldots, \beta^n\}$

Proof (Part 1)

- Let r be the number of positive values in $\{\alpha_1, \dots, \alpha_n\}$
- We can assume that

$$\alpha_k = Q(e_k, e_k) \begin{cases} > 0 & \text{if } 1 \leq k \leq r \\ \leq 0 & \text{if } r+1 \leq k \leq n \end{cases}$$

- Let R be the subspace spanned by $\{e_1, \dots, e_r\}$
- Similarly, let s be the number of positive values in $\{\beta_1, \dots, \beta_n\}$ and assume that

$$\beta_k = Q(f_k, f_k) \begin{cases} > 0 & \text{if } 1 \leq k \leq s \\ \leq 0 & \text{if } s+1 \leq k \leq n \end{cases}$$

- Let S be the subspace spanned by $\{f_1, \dots, f_s\}$

Proof (Part 2)

- Define the projection map

$$P : V \to R$$
$$e_1 v^1 + \cdots + e_n v^n \mapsto e_1 v^1 + \cdots + e_r v^r$$

- Let $P_S : S \to R$ be the restriction of P to S
- On one hand, if $v \in S$, then $v = f_1 b^1 + \cdots + f_s b^s$ and

$$Q(f_1 b^1 + \cdots + f_s b^2) = \beta_1 (b^1)^2 + \cdots + \beta_s (b^s)^2 \geq 0$$

- On the other hand, if $v \in \ker P_S$, then

$$v = e_{r+1} a^{r+1} + \cdots + e_n a^n$$

and therefore

$$Q(v, v) = \alpha_{r+1} (a^{r+1})^2 + \cdots + \alpha_n (a^n)^2 \leq 0$$

Proof (Part 3)

- It follows that

$$0 = Q(v, v) = \beta_1 (b^1)^2 + \cdots \beta_s (b^s)^2$$

and, since $\beta_1, \ldots, \beta_s \geq 0$,

$$\beta_1 = \cdots = \beta_s = 0$$

- Therefore, $\ker(P_S) = \{0\}$ and $s = \dim(S) \leq r = \dim(R)$
- The same argument with the bases switched implies that $r = \dim(R) \leq s = \dim(S)$
- The same argument proves that the number of negative values in $\{\alpha_1, \ldots, \alpha_n\}$ equals the number of negative values in $\{\beta_1, \ldots, \beta_n\}$
- It now follows that the number of zeros in $\{\alpha_1, \ldots, \alpha_n\}$ equals the number of zeros in $\{\beta_1, \ldots, \beta_n\}$
- Therefore, the signature of Q is well defined independent of the basis

Orthonormal Basis of a Quadratic Form

- Let $Q : V \to \mathbb{R}$ be a quadratic form with signature (p, q, r)
- There is a bilinear or sesquilinear form $B : V \times V \to \mathbb{F}$ such that
$$Q(v) = B(v, v)$$
- Then there exists a basis $(e_1, \dots, e_n)$ of V such that
$$B(e_j, e_k) = \begin{cases} 1 & \text{if } 1 \leq j = k \leq p \\ -1 & \text{if } p + 1 \leq j = k \leq p + q \\ 0 & \text{if } p + q + 1 \leq j = k \leq n \\ 0 & \text{if } j \neq k \end{cases}$$

Cayley-Hamilton Theorem

- Recall that the characteristic polynomial of a square matrix A is
$$p(x) = \det(A - xI)$$
- Given any polynomial
$$p(x) = a_0 + a_1 x + \cdots + a_n x^n,$$
and square matrix M, we can define
$$p(M) = a_0 I + a_1 M + \cdots + a_n M^n$$
- **Theorem:** If p is the characteristic polynomial of a square matrix A, then
$$p(M) = 0$$

Wrong Proof

- Since $p(x) = \det(A - xI)$,

$$p(A) = \det(A - AI) = 0$$

Characteristic Polynomial

- Recall that if A is a square polynomial over $\mathbb{C}$, its characteristic polynomial is

$$p_A(x) = \det(A - xI) = (\lambda_1 - x) \cdots (\lambda_n - x),$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, counting multiplicities

- Therefore, for each eigenvalue λ_k,

$$p_A(\lambda_k) = 0$$

Polynomial Function of Diagonal Matrix (Part 1)

- Given a polynomial

$$p(x) = a_0 + a_1 x + \cdots + a_k x^k,$$

and a diagonal matrix,

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

let

$$p(D) = a_0 I + a_1 D + \cdots + a_n D^n$$

Polynomial Function of Diagonal Matrix (Part 2)

- Therefore,

$$
\begin{aligned}
& p(D) \\
&= a_0 I + a_1 \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} + \cdots + a_n \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} \\
&= \begin{bmatrix} a_0 + a_1\lambda_1 + \cdots + a_n\lambda_1^n & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & a_0 + a_1\lambda_n + \cdots + a_n\lambda_n^n \end{bmatrix} \\
&= \begin{bmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{bmatrix}
\end{aligned}
$$

Proof of Cayley-Hamilton For Diagonal Matrix

- Therefore,

$$p_D(D) = \begin{bmatrix} p_D(\lambda_1) & 0 & \cdots & 0 \\ 0 & p_D(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & p_D(\lambda_n) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Cayley-Hamilton For Diagonalizable Matrix (Part 1)

- If $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A, then since

$$0 = p_A(\lambda_k) = \det(A - \lambda_k I)$$

- If A is diagonalizable, then there is an invertible matrix M such that

$$A = MDM^{-1},$$

where

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

Cayley-Hamilton For Diagonalizable Matrix (Part 2)

- Observe that for each positive integer k,

$$
\begin{aligned}
(MDM^{-1})^k &= (MDM^{-1})\cdots(MDM^{-1}) \\
&= MD(M^{-1}M)\cdots D(M^{-1}M)DM^{-1} \\
&= MD^kM^{-1}
\end{aligned}
$$

- Observe that

$$
\begin{aligned}
p_A(x) &= \det(A - xI) \\
&= \det(MDM^{-1} - M(xI)M^{-1}) \\
&= (\det(M))\det(D - xI)(\det(M^{-1}) \\
&= \det(D - xI) = p_D(x)
\end{aligned}
$$

- Therefore,

$$
\begin{aligned}
p_A(A) &= a_0I + a_1A + \cdots + a_nA^n \\
&= a_0MIM^{-1} + a_1MDM^{-1} + \cdots + a_n(MDM^{-1})^n \\
&= M(a_0I + a_1D + \cdots + a_nD^n)M^{-1} \\
&= Mp_D(D)M^{-1}
\end{aligned}
$$

Proof of Cayley-Hamilton Using Analysis

- For any square matrix A, there exists a sequence of diagonalizable matrices that converges to A
- The map

$$\begin{aligned} \mathrm{gl}(n, \mathbb{F}) \times \mathrm{gl}(n, \mathbb{F}) &\to \mathrm{gl}(n, \mathbb{F}) \\ (A, B) &\mapsto p_A(B) \end{aligned}$$

  is continuous
- Therefore,

$$p_A(A) = \lim_{k \to \infty} p_{A_k}(A_k) = 0$$