Multivariable Control

Lecture 02

Linear Algebrea

John T. Wen

September 1, 2006

Outline

- Linear spaces and mappings
- Matrix theory
- Singular value decomposition
- Vector and Matrix norms

Ref: Ch. 2 of Zhou/Doyle

Copyrighted by John T. Wen

Linear spaces and mappings

Linear space: $(\mathcal{V}, \mathbf{F}, +, \cdot)$, $\mathcal{V}$ is closed under vector addition and scalar multiplication (with $\mathbf{F}$).
Example: $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$ (over $\mathbb{R}$), $\mathbb{C}^n$, $\mathbb{C}^{n\times m}$ (over $\mathbb{C}$)

Subspace: subset of linear space that is also a linear space.

$\text{span}\{v_1,\ldots,v_n\}$, $v_i \in \mathcal{V}$: All linear combinations of v_i's, $\sum_i \alpha_i v_i$, $\alpha_i \in \mathbf{F}$.

Linear independence of $\{v_1,\ldots,v_n\}$, $v_i \in \mathcal{V}$: $\sum_i \alpha_i v_i = 0$ if and only if $\alpha_i = 0$. Matrix form: $[v_1 : \ldots : v_n]\,\alpha = 0$, $\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$, if and only if $\alpha = 0$.

Basis of $\mathcal{V}$: $\{v_1,\ldots,v_n\}$ that is linear independent and $\text{span}\{v_1,\ldots,v_n\} = \mathcal{V}$. Canonical basis of $\mathbb{R}^n$: unit vectors.

Dimension of $\mathcal{V}$: number of vectors in the basis set.

Copyrighted by John T. Wen

Linear Mapping

Linear mapping: $A : \mathcal{V} \to \mathcal{W}$ such that superposition holds, $A(\alpha v_1 + \beta v_2) = \alpha A v_1 + \beta A v_2$.

Representation of A as a matrix:
Choose $E_v = \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}$ as a basis of $\mathcal{V}$ and $E_w = \begin{bmatrix} w_1 & \ldots & w_m \end{bmatrix}$ as a basis of $\mathcal{W}$, then representation of A in these bases is $[A]$ (ith column is Av_i represented in E_w):

$$AE_v = E_w [A] .$$

Example: Let $\mathcal{V} = \text{span}\{0, 1, \sin t, \cos t\}$, $t \in \mathbb{R}_+$ ($t \geq 0$), and $A = d/dt$. Represent A in this basis.

Transformation between bases: $AE_{v_1} = E_{v_1} A_1$, $AE_{v_2} = E_{v_2} A_2$, then $T^{-1} A_1 T = A_2$, $T = E_{v_1}^{-1} E_{v_2}$.

Copyrighted by John T. Wen

Range and Null Spaces

Range space (or Image): $\mathcal{R}(A) := \left\{ w \in \mathcal{W} : \exists v \in \mathcal{V} \text{ such that } Av = w \right\}$. $\mathcal{R}(A)$ is a subspace of $\mathcal{W}$.

Rank: $\operatorname{rank}(A) = \dim(\mathcal{R}(A))$

A is *full rank* (or *surjective*, or *onto*) if $\mathcal{R}(A) = \mathcal{W}$

MATLAB: basis for range space `orth(A)`

Null space (or Kernel): $\mathcal{N}(A) := \left\{ v \in \mathcal{V} : \text{ such that } Av = 0 \right\}$. $\mathcal{N}(A)$ is a subspace of $\mathcal{V}$.

A is *injective*, or *one-to-one*, if $\mathcal{N}(A) = \{0\}$.

MATLAB: basis for null space `null(A)`

Consider $A : \mathcal{V} \to \mathcal{V}$. A-invariant subspace: $A\mathcal{S} \subset \mathcal{S}$. Choose another subspace $\mathcal{T}$ so that $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$. Representation of A is then $\begin{bmatrix} A_1 & A_2 \\ 0 & A_4 \end{bmatrix}$.

What will $[A]$ look like if we order the basis so that $\mathcal{V} = \mathcal{T} \oplus \mathcal{S}$?

What will $[A]$ look like if $\mathcal{T}$ is also A-invariant?

Copyrighted by John T. Wen

Matrix Theory

Now consider a linear mapping $A : \mathcal{V} \to \mathcal{V}$, $\dim(\mathcal{V}) = n$, represented in a basis, $[A]$ ($n \times n$ square matrix). From now on, we'll just write A for $[A]$.

Jordan Decomposition

Consider $A : \mathbb{C}^n \to \mathbb{C}^n$. $Ax = \lambda x$, $x =$ eigenvector, $\lambda =$ eigenvalue. Then $\mathcal{N}(\lambda I - A)$ is A-invariant.

If there are n independent eigenvectors, they form a basis; A represented in this basis is diagonal: $T^{-1}AT = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ where $T = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}$.

Copyrighted by John T. Wen

Jordan Decomposition

In general, A can be transformed to a block diagonal form: $T^{-1}AT = \text{diag}\{J_1, \ldots, J_p\}$, J_i is called a Jordan block. There are two types of Jordan blocks:

- Simple: $J_i = \lambda_i$
- Degenerate: $Ax_i^{(1)} = \lambda_i x_i^{(1)}$, $Ax_i^{(2)} = \lambda_i x_i^{(2)} + x_i^{(1)}$, ..., $Ax_i^{(k_i)} = \lambda_i x_i^{(k_i)} + x_i^{(k_{i-1})}$. $x_i^{(\ell)}$ are called generalized eigenvectors. $J_i = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \lambda & \ddots & \\ & & & 1 \\ 0 & & 0 & \lambda_i \end{bmatrix}$

If we consider $A : \mathbb{R}^n \to \mathbb{R}^n$, there is one more type of block when $\lambda = \alpha + \beta j$. In this case

$$Ax_r = \alpha x_r - \beta x_i, \quad Ax_i = \alpha x_i + \beta x_r,$$

so the corresponding Jordan block is

$$\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}.$$

Copyrighted by John T. Wen

Special Classes of Matrices

Self-adjoint matrix (symmetric or Hermitian): $A = A^*$. All eigenvalues are real and all eigenvectors form an orthonormal basis (A is diagonalizable, $A = U\Lambda U^*$, U is unitary).

Unitary matrix: $A^* = A^{-1}$ (columns, or rows, are orthonormal).

Positive (semi) definite ($A > 0$ or $A \geq 0$): $x^T Ax \geq \delta \|x\|^2$, $\delta > (\geq) 0$. W.l.o.g., A is symmetric (all eigenvalues are positive or non-negative). If $A \geq 0$ and of rank r, then $\exists B$ of rank r such that $A = BB^T$.

Copyrighted by John T. Wen

Matrix Inversion Formula

Let $A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, A_{11}, A_{22} square.

If A_{11} nonsingular, then

$$A = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}.$$

where $\Delta := A_{22} - A_{21}A_{11}^{-1}A_{12}$. A is nonsingular if and only if Δ is nonsingular.
If A_{22} nonsingular, then

$$A = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{\Delta} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix}.$$

where $\hat{\Delta} := A_{11} - A_{12}A_{22}^{-1}A_{21}$. A is nonsingular if and only if $\hat{\Delta}$ is nonsingular.

Schur complement of A: Δ (or $\hat{\Delta}$).

Copyrighted by John T. Wen

Singular Value Decomposition

Given $A \in \mathbb{C}^{m\times n}$. There exist unitary matrices $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$, such that

$$A = U\Sigma V^*$$

where $\Sigma = \begin{bmatrix} \operatorname{diag}\{\sigma_1, \ldots, \sigma_r\} & 0 \\ 0 & 0 \end{bmatrix}$ with $r = \operatorname{rank}(A)$.

SVD can be visualized as mapping between unit ball in $\mathbb{C}^n$ to ellipsoid in $\mathbb{C}^m$ with principal axes given by u_i's and lengths given by σ_i.

Note that non-zero singular values are the square root of non-zero eigenvalues of A^*A and AA^*.

Copyrighted by John T. Wen

Useful Inequalities for Singular Values

Define $\overline{\sigma}(A)$ ($\underline{\sigma}(A)$) as the maximum (minimum) singular value of A.

1. $|\underline{\sigma}(A+\Delta)-\underline{\sigma}(A)| \leq \overline{\sigma}(\Delta)$
2. $\underline{\sigma}(A\Delta) \geq \underline{\sigma}(A)\underline{\sigma}(\Delta)$
3. $\overline{\sigma}(A^{-1}) = \frac{1}{\underline{\sigma}(A)}$ if A is invertible

Copyrighted by John T. Wen

Vector Norms

How do we characterize the "size" of a vector?

Def: Normed Linear Space: a vector space $\mathcal{V}$ endowed with a norm, $\|\cdot\|_{\mathcal{V}} : \mathcal{V} \to [0,\infty)$, which satisfies
(a) $\|v\|_{\mathcal{V}} \geq 0$ (positivity) (b) $\|v\|_{\mathcal{V}} = 0$ if and only if $v = 0$ (positive definiteness)
(c) $\|\alpha v\|_{\mathcal{V}} = |\alpha| \, \|v\|_{\mathcal{V}}$ for all $v \in \mathcal{V}$ (homogeneity)
(d) $\|u+v\|_{\mathcal{V}} \leq \|u\|_{\mathcal{V}} + \|v\|_{\mathcal{V}}$, for all $u, v \in \mathcal{V}$ (triangular inequality)

If $x \in \mathbb{C}^n$, $\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, $1 \leq p \leq \infty$. Most common ones: $p = 1, 2, \infty$. 2-norm is also called Euclidean norm.

If $n = 2$, what does $\|x\|_p = 1$ (unit ball in p-norm) look like?

Copyrighted by John T. Wen

Inner Product and Orthogonality

The concept of projection and orthogonality requires additional algebraic structure called inner product (like dot product for vectors).

Def: Inner product space: a vector space $\mathcal{V}$ endowed with an inner product, $< \cdot, \cdot >_{\mathcal{V}}$: $\mathcal{V} \times \mathcal{V} \to \mathbf{F}$ ($\mathbb{R}$ or $\mathbb{C}$) which satisfies

(a) $< v, v >_{\mathcal{V}} \geq 0$ for all $v \in \mathcal{V}$;

(b) $< v, v >_{\mathcal{V}} = 0$ if and only if $v = 0$;

(c) $< v, \alpha_1 u_1 + \alpha_2 u_2 >_{\mathcal{V}} = \alpha_1 < v, u_1 >_{\mathcal{V}} + \alpha_2 < v, u_2 >_{\mathcal{V}}$;

(d) $< u, v >$ is the complex conjugate of $< v, u >$.

Inner product space (IPS) is a special case of normed linear space (NLS) since $\sqrt{< v, v >}$ is a norm. An NLS is an IPS if and only if $\|\cdot\|$ satisfies the parallelogram law: $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

Orthogonality: $u \perp v$ means $< u, v >= 0$ (nec. & suf. cond: $\|u+v\|^2 = \|u\|^2 + \|v\|^2$).

Cauchy-Schwarz Inequality: $|< u, v >| \leq \|u\| \, \|v\|$

Example: $\mathbb{R}^n$: $< x, y >= x^T y$, $\mathbb{C}^n$: $< x, y >= x^* y$ (among all p-norm's, only 2-norm is induced by an inner product).

Copyrighted by John T. Wen

Matrix Norms

Consider $A \in \mathbb{C}^{p \times m}$. How to we characterize the size of A?

Vector Approach: Regard A as an $p \cdot m$ vector (e.g., by stacking up the columns) and apply any of the vector norm.

Examples:

If vector 2-norm is used, then $\|A\| = \sqrt{tr(A^*A)}$ which is defined as the Frobenius norm, $\|A\|_F$.

If vector ∞-norm is used, then $\|A\| = \max\left\{\left|A_{ij}\right|\right\}$ which is called the max norm, $\|A\|_{\max}$.

If vector 1-norm is used, then $\|A\| = \sum_{i,j}\left\{\left|A_{ij}\right|\right\}$.

Copyrighted by John T. Wen

Matrix Norms (cont.)

Operator Approach: Regard A as a mapping of $\mathbb{C}^m$ to $\mathbb{C}^p$. Induced norm:

$$\|A\|_{i,p} = \sup_{x \in \mathbb{C}^m} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p = 1} \|Ax\|_p \, .$$

Example:

$\|A\|_{i,2} = \overline{\sigma}(A)$ (most common, default norm)

$\|A\|_{i,1} = \max_j \sum_i |A_{ij}|$ = maximum column sum

$\|A\|_{i,\infty} = \max_i \sum_j |A_{ij}|$ = maximum row sum

Submultiplicative property: $\|AB\| \leq \|A\| \, \|B\|$. True for any induced norm, but may not hold for other matrix norms.

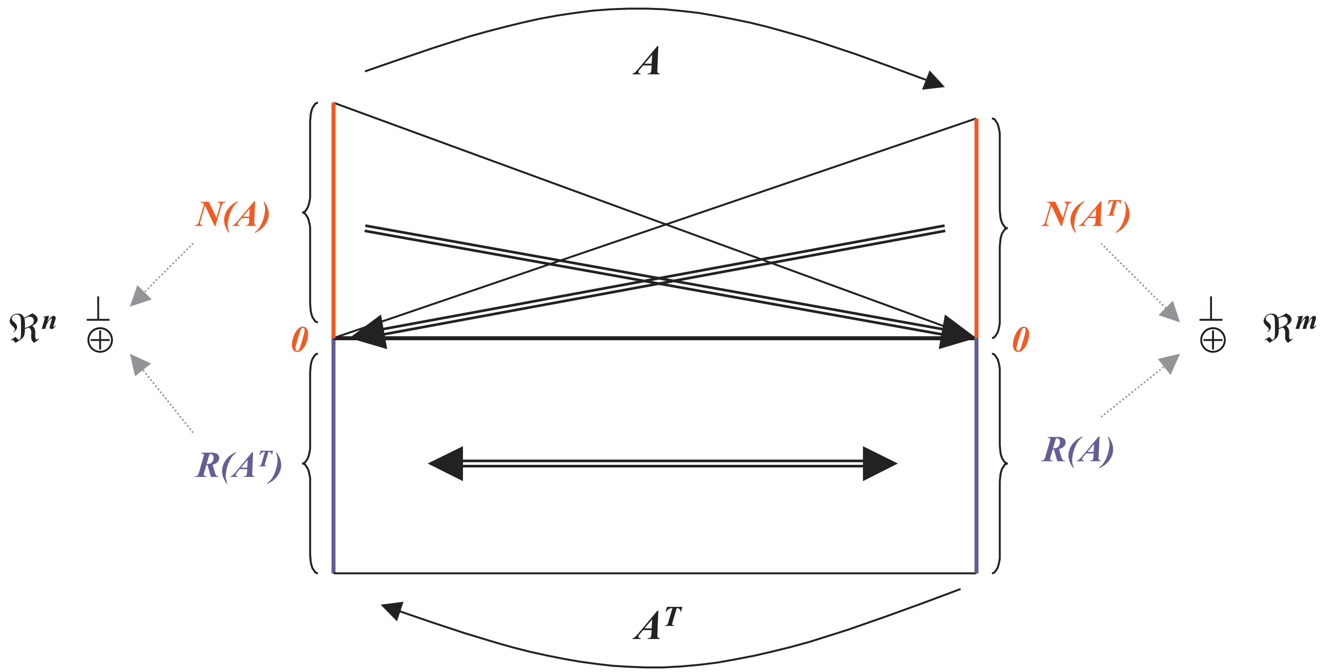
Copyrighted by John T. Wen

Some Useful Matrix Inequalities

1. $\rho(A) \leq \|A\|$ (for any induced norm and F-norm), $\rho(A) := \max_i |\lambda_i(A)|$ is the spectral radius.
2. $\|AB\| \leq \|A\| \, \|B\|$ (submultiplicative property; true for any induced norm).
3. $\|UAV\| = \|A\|$ for any unitary matrices U, V (true for any induced norm and F-norm).
4. $\|AB\|_F \leq \|A\| \, \|B\|_F \leq \|B\| \, \|A\|_F$ (for induced-2-norm).

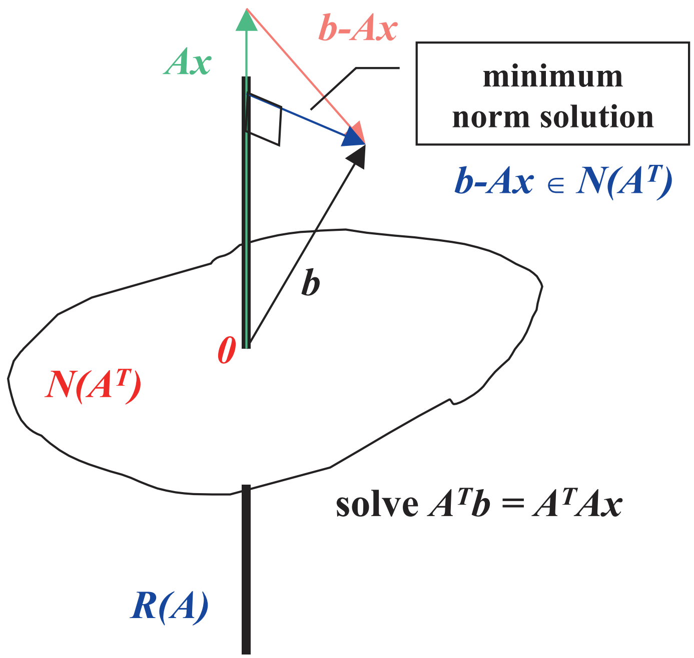
Copyrighted by John T. Wen

Orthogonal Decomposition



Copyrighted by John T. Wen

Least Square Error Solution

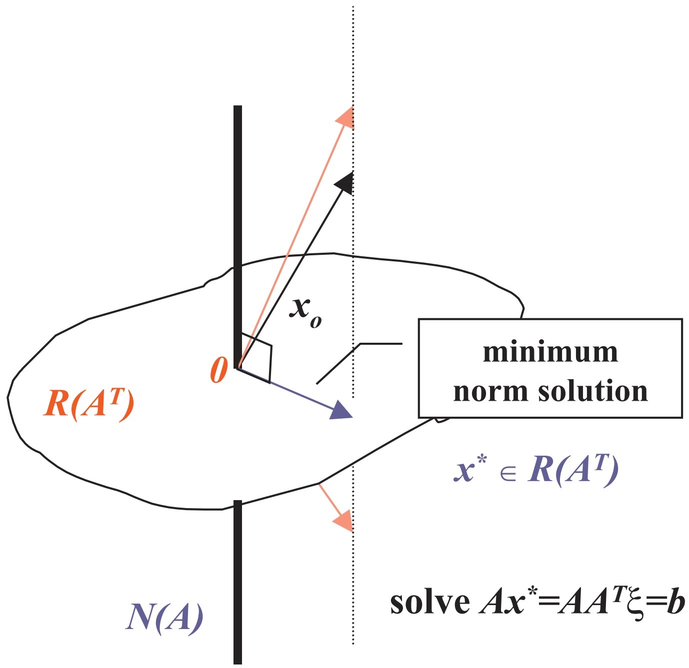


If A is full (row) rank $\xi = (AA^T)^{-1}b$ and $x^* = A^T\xi = A^T(AA^T)^{-1}b$.

Copyrighted by John T. Wen

Least Square Norm Solution

Minimum Norm Solution



If A is full (column) rank $x^* = (A^T A)^{-1} A^T b$.

Copyrighted by John T. Wen

Generalized Inverse

- Given $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{n\times m}$ is a right inverse if $AX = I_m$, $Y \in \mathbb{R}^{n\times m}$ is a left inverse if $YA = I_n$. If A is of full row rank, *one* of the right inverse is $X = A^T(AA^T)^{-1}$. If A is of full column rank, *one* of the left inverse is $Y = (A^TA)^{-1}A^T$.
- Pseudo-inverse (or Moore-Penrose inverse) A^+ (uniquely defined):
  (i) $AA^+A = A$ (ii) $A^+AA^+ = A^+$ (iii) $(AA^+)^T = AA^+$ (iv) $(A^+A)^T = A^+A$
  SVD for A: $A = U\Sigma V^T$, $\Sigma = \text{diag}\{\Sigma_r, 0\}$, then $A^+ = V\Sigma^+U^T$, $\Sigma^+ = \text{diag}\{\Sigma_r^{-1}, 0\}$.
  Decompose $A = BC$, B of full column rank, C of full row rank. Then
  $$A^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T.$$
  (Verify that the four conditions are satisfied.)

Copyrighted by John T. Wen