Systems of linear equations

- An arbitrary system of m linear equations in n unknowns can be written as:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

or $\underline{A}\vec{x} = \vec{b}$

where: $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

Example (m=n=2):

$$\begin{aligned} 2x_1 + 3x_2 &= 8 \\ 4x_1 + 6x_2 &= 9 \end{aligned} \Rightarrow \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \end{bmatrix}$$

63

$\underline{A}\vec{x} = \vec{b}$

If $\vec{b} \in Col(\underline{A})$

$\Rightarrow$ solution exists

$$[\vec{c}_1 \ \vec{c}_2 \ \ldots \ \vec{c}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{c}_1 + x_2\vec{c}_2 + \ldots + x_n\vec{c}_n$$

64

Solution is unique

when rank($\underline{A}$) = n

$\Rightarrow$ solution is unique

nullity = 0

65

Solutions of Ax=b (m = n)

- Characterize the solutions of Ax=b using conditions on the rank of A and A|b (i.e., augmented matrix).

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \begin{matrix} b_1 \\ b_2 \\ \\ b_m \end{matrix}$$

66

Solutions of Ax=b (m=n)

(2) The system has *no solution* if

rank(A|b) > rank(A)

Consider the system,

$$\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \end{bmatrix} \quad or \quad \begin{aligned} 2x_1 + 3x_2 &= 8 \\ 4x_1 + 6x_2 &= 9 \end{aligned}$$

$$\begin{vmatrix} 2 & 3 \\ 4 & 6 \end{vmatrix} = 0 \Rightarrow \text{singular}$$

$$rank \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} = 1$$

$$rank \begin{bmatrix} 2 & 3 & 8 \\ 4 & 6 & 9 \end{bmatrix} = 2 \Rightarrow rank(A\,|\,b) > rank(A)$$

b cannot be expressed as a linear combination of the columns of A

e.g., using substitution leads to the contradiction 16=9

67

Solutions of Ax=b (m=n)

(1) The system has *one solution* if :

rank(A|b) = rank(A) = n

Solution: $x = A^{-1}b$

i.e., b be expressed as a linear combination of the columns of A:

68

Solutions of Ax=b (m=n)

(3) The system has *infinitely many solutions* if

$rank(A|b) = rank(A) < n$

- Less equations than unknowns (i.e, free variables).

- b can be expressed as a linear combination of the columns of A in more than one ways.

69

Solutions of Ax=b (m=n)

The following statements are equivalent:

(a) $rank(A|b) = rank(A) = n$

(b) A is invertible

(c) $det(A) \neq 0$

(d) b has a unique expansion in the column space of A

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix} x_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{bmatrix} x_2 + \cdots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{bmatrix} x_n = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix}$$

70

Homogeneous system: Ax=0 (m = n)

- If b=0, then Ax=0 is called homogeneous.

(1) Has the trivial solution x=0

iff $rank(A) = n$ (i.e., A is invertible)

(2) Has a non-trivial solution

iff $rank(A) < n$ (i.e., A is singular)

(3) if rank (A) = n-1 → unique non-trivial solution

71

Over/Under determined Systems

When m>n the system is called *over-determined*.

When m<n the system is called *under-determined*.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

72

Solving Ax=b ($m > n$)

- Consider the *over-determined* system of linear equations: $Ax = b,\ (A \text{ is } m\text{x}n \text{ with } m>n)$
- Let r be the residual vector for some x:

$$\vec{r} = Ax - b$$

$\vec{b}$ observation

Ax model estimate

- The vector x^* which yields the smallest possible residual is called a *least-squares* solution:

$$||r|| = ||Ax^* - b|| \leq ||Ax - b|| \text{ for all } x \in R^n$$

73

Solving Ax=b ($m > n$) (cont'd)

- Although a least-squares solution always exist, it might not be unique!

$A^T \; A_{m\text{x}n} \vec{x} = A^T b$

- The least-squares solution x with the smallest norm $||x||$ is unique and it is given by:

$n\text{x}n$

$$Ax = b \Rightarrow A^T A x = A^T b \Rightarrow x = (A^T A)^{-1} A^T b = A^+ b$$

$A^+ = (A^T A)^{-1} A^T$ ← pseudoinverse

74

Solving Ax=b ($m > n$) - Example

$-11x_1 + 2x_2 = 0$
$2x_1 + 3x_2 = 7$
$2x_1 - x_2 = 5$

pinv(A)

$$\begin{bmatrix} -11 & 2 \\ 2 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 5 \end{bmatrix}$$

$$x = A^{+}b = \begin{bmatrix} -.148 & .180 & .246 \\ .164 & .189 & -.107 \end{bmatrix} \begin{bmatrix} 0 \\ 7 \\ 5 \end{bmatrix} = \begin{bmatrix} 2.492 \\ 0.787 \end{bmatrix}$$

$(A^T A)^{-1} A^T$

75

Homogeneous systems

$$Ax = 0$$

- The minimum-norm solution is x=0; need to modify the meaning of a least-squares solution by imposing the constraint:

$$\|x\| = 1$$

- This is a "constrained" optimization problem:

$$\min_{\|x\|=1} \|Ax\|$$

$\|A\vec{x} - \vec{b}\|$

residual

76

SVD (Singular Value Decomposition)

- Any real ***m* x *n*** matrix *A* can be decomposed uniquely:

$$A = UDV^T$$

- *U* is ***m* x *n*** and column orthonormal (U^TU=I)
- *D* is ***n* x *n*** and diagonal

$$D = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

  - σ_i are called *singular* values of A
  - It is assumed that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$
- *V* is ***n* x *n*** and orthonormal (VV^T=V^TV=I)

77

$A_{m \times n}$ → does not have eigenvalue decomposition

$\underset{n \times n}{B} = \underset{n \times m}{A^T} \underset{m \times n}{A}$ → $B^T = (A^T A)^T = A^T (A^T)^T = A^T A = B$

Symmetric

$= V \Lambda_B V^T$

V rotation matrix

$V^{-1} = V^T$ $\det V = 1$ always diagonalized

78

$\underset{m \times m}{C} = \underset{m \times n}{A} \underset{n \times m}{A^T}$

$= U \Lambda_2 U^T$

$U^{-1} = U^T$ rotation matrix

its columns are eigenvectors of AA^T

79

Remark non-zero eigenvalues of A^TA and AA^T are the same.

Remark eigenvalues of A^TA and AA^T are nonnegative ($\lambda_i \geq 0$)

80

SVD (cont'd)

- If ***m*=*n*,** then:

$$A = UDV^T$$

- U is ***n* x *n*** and orthonormal ($U^TU=UU^T=I$)
- D is ***n* x *n*** and diagonal

$$D = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$$

- V is ***n* x *n*** and orthonormal ($VV^T=V^TV=I$)

81

SVD (cont'd)

- The columns of U are eigenvectors of AA^T

for square matrices: $A=P\Lambda P^{-1}$

$$AA^T = UDV^TVDU^T = UD^2U^T$$

- The columns of V are eigenvectors of A^TA

$$A^TA = VDU^TUDV^T = VD^2V^T$$

- If λ_i is an eigenvalue of A^TA (or AA^T), then $\lambda_i = \sigma_i^2$

82

SVD - Example

$$A = \begin{bmatrix} 0.9501 & 0.8913 & 0.8214 & 0.9218 \\ 0.2311 & 0.7621 & 0.4447 & 0.7382 \\ 0.6068 & 0.4565 & 0.6154 & 0.1763 \\ 0.4860 & 0.0185 & 0.7919 & 0.4057 \end{bmatrix} \Rightarrow A = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

$U = (u_1\ u_2\ \ldots\ u_n)$ $\qquad$ $V = (v_1\ v_2\ \ldots\ v_n)$

$$U = \begin{bmatrix} 0.7301 & 0.1242 & 0.1899 & -0.6445 \\ 0.4413 & 0.6334 & -0.3788 & 0.5104 \\ 0.3809 & -0.3254 & 0.6577 & 0.5626 \\ 0.3564 & -0.6910 & -0.6229 & 0.0871 \end{bmatrix} \quad V = \begin{bmatrix} 0.4903 & -0.4004 & 0.5191 & -0.5743 \\ 0.4770 & 0.6433 & 0.4642 & 0.3783 \\ 0.5362 & -0.5417 & -0.2770 & 0.5850 \\ 0.4945 & 0.3638 & -0.6620 & -0.4299 \end{bmatrix}$$

$$D = \begin{bmatrix} 2.4479 & 0 & 0 & 0 \\ 0 & 0.6716 & 0 & 0 \\ 0 & 0 & 0.3646 & 0 \\ 0 & 0 & 0 & 0.1927 \end{bmatrix}$$

83

SVD – Another Example

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix},$$

The eigenvalues of AA^T, A^TA are: $\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 28.86 \\ 0.14 \\ 0 \end{bmatrix}$

The eigenvectors of AA^T, A^TA are:

$$u_1 = v_1 = \begin{bmatrix} 0.454 \\ 0.766 \\ 0.454 \end{bmatrix}, u_2 = v_2 = \begin{bmatrix} 0.542 \\ -0.643 \\ 0.542 \end{bmatrix}, u_3 = v_3 = \begin{bmatrix} -0.707 \\ 0 \\ -0.707 \end{bmatrix}$$

84

SVD properties

- A square (n × n) matrix A is singular iff at least one of its singular values $\sigma_1, \ldots, \sigma_n$ is zero.
- The rank of matrix A is equal to the number of nonzero singular values σ_i

85

Computing A^{-1} using SVD

- If A is a n x n nonsingular matrix, then its inverse can be computed as follows:

$$A = UDV^T \Rightarrow A^{-1} = VD^{-1}U^T$$

$$A^{-1} = V\left[diag(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1})\right]U^T$$ easy to compute!

($U^TU=UU^T=I$ or $U^T=U^{-1}$ and $V^TV=VV^T=I$ or $V^T=V^{-1}$)

86

Least–squares Solution of Homogeneous Equations

87

Derivation II — SVD

- Let $A = USV^\top$, where U is $m \times n$ orthonormal, S is $n \times n$ diagonal with descending order, and $V^\top$ is $n \times n$ also orthonormal.
- From orthonormality of U, V follows that $\|USV^\top \mathbf{h}\| = \|SV^\top \mathbf{h}\|$ and $\|V^\top \mathbf{h}\| = \|\mathbf{h}\|$.
- Substitute $\mathbf{y} = V^\top \mathbf{h}$. Now, we minimize $\|S\mathbf{y}\|$ subject to $\|\mathbf{y}\| = 1$.
- Remember that S is diagonal and the elements are sorted descendently. Than, it is clear that $\mathbf{y} = [0, 0, \ldots, 1]^\top$.
- From substitution we know that $\mathbf{h} = V\mathbf{y}$ from which follows that sought $\mathbf{h}$ is the last column of the matrix V.

88

- To find an extreme (the sought $\mathbf{h}$) we must solve $\frac{\partial}{\partial \mathbf{h}} (\mathbf{h}^\top A^\top A \mathbf{h} + \lambda(1 - \mathbf{h}^\top \mathbf{h})) = 0$.
- We derive: $2A^\top A\mathbf{h} - 2\lambda\mathbf{h} = 0$.
- After some manipulation we end up with: $(A^\top A - \lambda E)\mathbf{h} = 0$ which is the characteristic equation. Hence, we know that $\mathbf{h}$ is an eigenvector of $(A^\top A)$ and λ is an eigenvalue.
- The least-squares error is $e = \mathbf{h}^\top A^\top A \mathbf{h} = \mathbf{h}^\top \lambda \mathbf{h}$.
- The error will be minimal for $\lambda = \min_i \lambda_i$ and the sought solution is then the eigenvector of the matrix $(A^\top A)$ corresponding to the smallest eigenvalue.

89

Homogeneous systems (cont'd)

- The $\min_{\|x\|=1} \|Ax\|$ solution for homogeneous systems is not always unique.
- Special case: $rank(A) = n - 1$ ($m \geq n - 1$, σ_n=0)

Solution: $x = av_n$ (a is a constant)

(v_n is the last column of V; the one corresponding to the smallest σ) $A = UDV^T$

90