Solving linear systems and least square fit

Kumar Appaiah & Mahesh B. Patil & Madhu N. Belur

Department of Electrical Engg, IIT Bombay

http://www.ee.iitb.ac.in/%7Ebelur/talks/pdfs/cep19feb.pdf

CEP, Feb 2019

Simultaneous system of equations

Solve for x_1 and x_2

$$x_1 + x_2 = 9$$

- Line in a plane. Plane: set of all x_1 and x_2 points: $\mathbb{R}^2$
- Multiple solutions
- 1 equation, 2 unknowns: under-determined

Simultaneous system of equations

Solve for x_1 and x_2

$$\begin{aligned} x_1 + x_2 &= 9 \\ 2x_1 + 3x_2 &= 19 \end{aligned}$$

- 2 equation, 2 unknowns: determined (if $\cdots$)

Simultaneous system of equations

Solve for x_1 and x_2

$$\begin{aligned} x_1 + x_2 &= 9 \\ 2x_1 + 3x_2 &= 19 \end{aligned}$$

- 2 equation, 2 unknowns: determined (if $\cdots$)
- Two lines in a plane
- (Hopefully) intersect at one point
- Instead: parallel lines (no solution)
- Or same line (multiple solutions)

Over-determined system of equations

$$\begin{aligned} x_1 + x_2 &= 9 \\ 2x_1 + 3x_2 &= 19 \\ 2x_1 + 5x_2 &= 10 \end{aligned}$$

- 3 equations, 2 unknowns: over-determined

Over-determined system of equations

$$\begin{aligned} x_1 + x_2 &= 9 \\ 2x_1 + 3x_2 &= 19 \\ 2x_1 + 5x_2 &= 10 \end{aligned}$$

- 3 equations, 2 unknowns: over-determined (insufficient ‘degrees’ of freedom)
- Unless some equations are ‘repeat’, maybe no exact solution
- We can aim for ‘nearest’ solution, i.e. find x_1 and x_2 to minimize

$$(x_1 + x_2 - 9)^2 + (2x_1 + 3x_2 - 19)^2 + (2x_1 + 5x_2 - 10)^2$$

(Least square fit)

More generality

Given $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$, solve for $x \in \mathbb{R}^n$ such that $Ax = b$.

- Rank: number of independent rows (same as number of independent columns): redundancy of equations taken care
- Under-determined : rank $A < n$
- Determined : rank $A = n$
- Over-determined : rank $A > n$

More generality

Given $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$, solve for $x \in \mathbb{R}^n$ such that $Ax = b$.

- Rank: number of independent rows (same as number of independent columns): redundancy of equations taken care
- Under-determined : rank $A < n$
- Determined : rank $A = n$
- Over-determined : rank $A > n$

Matrix 'language':

$$\begin{aligned} x_1 + x_2 &= 9 \\ 2x_1 + 3x_2 &= 19 \\ 2x_1 + 5x_2 &= 10 \end{aligned} \qquad \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 19 \\ 10 \end{bmatrix}$$

Solving circuits, KCL/KVL, (differential)-equations, discretized system of equations, simulation, $\cdots$

Inner product/outer-product way of understanding

Note: $Ax = \begin{bmatrix}1\\2\\2\end{bmatrix} x_1 + \begin{bmatrix}1\\3\\5\end{bmatrix} x_2$ and inner-product $\begin{bmatrix}7 & -2 & 2\end{bmatrix} \begin{bmatrix}1\\2\\5\end{bmatrix} = 13$

Inner product/outer-product way of understanding

Note: $Ax = \begin{bmatrix}1\\2\\2\end{bmatrix} x_1 + \begin{bmatrix}1\\3\\5\end{bmatrix} x_2$ and inner-product $\begin{bmatrix}7 & -2 & 2\end{bmatrix} \begin{bmatrix}1\\2\\5\end{bmatrix} = 13$

Outer-product $\begin{bmatrix}1\\2\\5\end{bmatrix} \begin{bmatrix}7 & -2 & 2\end{bmatrix} = \begin{bmatrix}7 & -2 & 2\\14 & -4 & 4\\35 & -10 & 10\end{bmatrix}$

Inner product/outer-product way of understanding

Note: $Ax = \begin{bmatrix}1\\2\\2\end{bmatrix} x_1 + \begin{bmatrix}1\\3\\5\end{bmatrix} x_2$ and inner-product $\begin{bmatrix}7 & -2 & 2\end{bmatrix}\begin{bmatrix}1\\2\\5\end{bmatrix} = 13$

Outer-product $\begin{bmatrix}1\\2\\5\end{bmatrix}\begin{bmatrix}7 & -2 & 2\end{bmatrix} = \begin{bmatrix}7 & -2 & 2\\14 & -4 & 4\\35 & -10 & 10\end{bmatrix}$

Note: Matrix-matrix product: both inner-product and outer-product viewpoints work.

Recommended to use both viewpoints everywhere

More of matrix language and manipulations

(Under any sequence of) following operations) solution is unchanged

- Interchange two rows (equations)
- Scaling a row by nonzero number
- Adding an equation to another

Above is same as pre-multiplying A by a nonsingular matrix U (and same to b too).

For numerical accuracy purposes, we will deal with more special nonsingular matrices:

Orthogonal matrix U: special square, nonsingular matrix: $UU^T = I$ (identity matrix)

U^T: Transpose of U.

SVD: Singular Value Decomposition

Given $A \in \mathbb{R}^{m \times n}$, there exist orthogonal matrices U and V and a 'diagonal' matrix Σ such that
$A = U\Sigma V^T$.
The matrix $\Sigma \in \mathbb{R}^{m \times n}$ has along the diagonals $\sigma_1, \sigma_2, \cdots, \sigma_{\min(m,n)}$. Further, one can ensure that $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_{\min(m,n)} \geqslant 0$. Once this is ensured, $\sigma_1, \sigma_2, \ldots, \sigma_{\min(m,n)}$ are called the singular values. They are unique and depend only on A, though U and V might be nonunique.
Note: because Σ is diagonal, $U\Sigma V^T = A$ can be viewed as expressing A as a sum of various rank-one matrices: each rank one matrix is an outer-product of two vectors u_i and v_i^T (and scaled by the singular value):
these two vectors u_i and v_i are corresponding columns of orthogonal matrices U and V.

SVD (contd)

Following are equivalent.

1. A has rank r.
2. Exactly r of the singular values $\sigma_1, \ldots, \sigma_r$ are nonzero and positive. Others are zero.
3. When attempting to express A as a sum of p number of rank-1 matrices, then the smallest such p is r.
4. Dimension of image of A is r, i.e. there are r number of independent vectors in the image of A. (Image A means range-space of A, i.e. column-span of A.)
5. r satisfies $n = r+$ dimension of nullspace of A.
   Nullspace of A is also called kernel of A. This is set of all vectors x such that $Ax = 0$.
6. When looking for $p \times p$ nonsingular submatrix within matrix A, the largest p such that there is a nonsingular submatrix is $p = r$.

Links for python codes used in the lecture:
http://www.ee.iitb.ac.in/%7Ebelur/talks/pdfs/cep19svd.py
http://www.ee.iitb.ac.in/%7Ebelur/talks/pdfs/cep19leastsq.py