

Interpreting Linear Equations

Stephen Boyd and Sanjay Lall
Revision by Amirhossein Afsharrad

EE263
Stanford University

Broad categories of applications

linear model or function $y = Ax$

some broad categories of applications:

- ▶ estimation or inversion
- ▶ control or design
- ▶ mapping or transformation

(this list is not exclusive; can have combinations ...)

Estimation or inversion

$$y = Ax$$

- ▶ y_i is i th measurement or sensor reading (which we know)
- ▶ x_j is j th parameter to be estimated or determined
- ▶ a_{ij} is sensitivity of i th sensor to j th parameter

sample problems:

- ▶ find x , given y
- ▶ find all x 's that result in y (*i.e.*, all x 's consistent with measurements)
- ▶ if there is no x such that $y = Ax$, find x s.t. $y \approx Ax$ (*i.e.*, if the sensor readings are inconsistent, find x which is almost consistent)

Control or design

$$y = Ax$$

- ▶ x is vector of design parameters or inputs (which we can choose)
- ▶ y is vector of results, or outcomes
- ▶ A describes how input choices affect results

sample problems:

- ▶ find x so that $y = y_{\text{des}}$
- ▶ find all x 's that result in $y = y_{\text{des}}$ (*i.e.*, find all designs that meet specifications)
- ▶ among x 's that satisfy $y = y_{\text{des}}$, find a small one (*i.e.*, find a small or efficient x that meets specifications)

Mapping or transformation

- ▶ x is mapped or transformed to y by linear function $y = Ax$

sample problems:

- ▶ determine if there is an x that maps to a given y
- ▶ (if possible) find *an* x that maps to y
- ▶ find *all* x 's that map to a given y
- ▶ if there is only one x that maps to y , find it (*i.e.*, decode or undo the mapping)

Matrix multiplication as mixture of columns

write $A \in \mathbb{R}^{m \times n}$ in terms of its columns

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

where $a_j \in \mathbb{R}^m$. Then then $y = Ax$ means

$$y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

(x_j 's are scalars, a_j 's are m -vectors)

- ▶ y is a (linear) combination or mixture of the columns of A
- ▶ coefficients of x give coefficients of mixture
- ▶ each column of A represents an *actuator*

Geometric interpretation of control

example: $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$, $y = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$

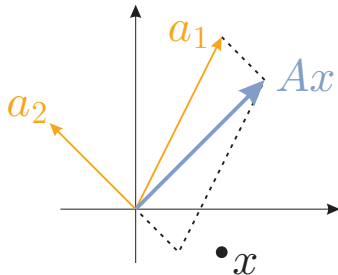
$$Ax = a_1 + (-0.5)a_2 = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$$

another example:

$$a_j = Ae_j$$

where e_j is the j th unit vector:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$



► j th column of A gives response to unit j th input

Matrix multiplication as inner product with rows

write A in terms of its rows:

$$A = \begin{bmatrix} \tilde{a}_1^\top \\ \tilde{a}_2^\top \\ \vdots \\ \tilde{a}_m^\top \end{bmatrix}$$

where $\tilde{a}_i \in \mathbb{R}^n$

then $y = Ax$ can be written as

$$y = \begin{bmatrix} \tilde{a}_1^\top x \\ \tilde{a}_2^\top x \\ \vdots \\ \tilde{a}_m^\top x \end{bmatrix}$$

- ▶ $y_i = \tilde{a}_i^\top x$, so that y_i is inner product of i th row of A with x
- ▶ each row of A represents a *sensor*

Geometric interpretation of estimation

$$a_i^\top x = \text{constant}$$

is a (hyper-)plane in $\mathbb{R}^n$ normal to a_i .

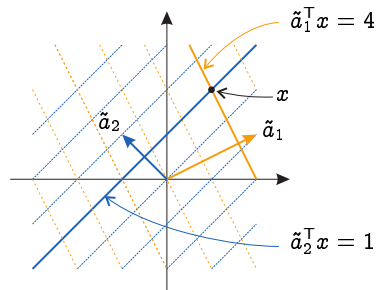
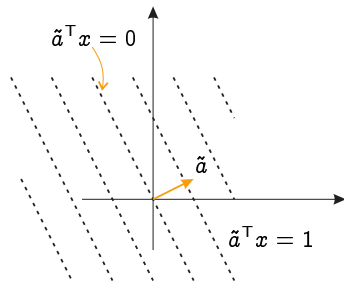
if $Ax = y$ then x is on intersection of hyperplanes

$$a_i^\top x = y_i$$

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

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$y = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$



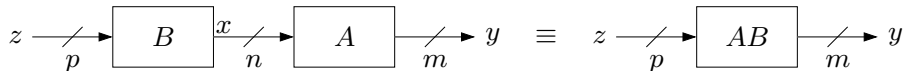
Matrix multiplication as composition

for $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$, $C = AB \in \mathbb{R}^{m \times p}$ where

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

composition interpretation

$y = Cz$ represents composition of $y = Ax$ and $x = Bz$



(note that B is on left in block diagram)

Column and row interpretations

can write product $C = AB$ as

$$C = \begin{bmatrix} c_1 & \cdots & c_p \end{bmatrix} = AB = \begin{bmatrix} Ab_1 & \cdots & Ab_p \end{bmatrix}$$

i.e., i th column of C is A acting on i th column of B

similarly we can write

$$C = \begin{bmatrix} \tilde{c}_1^\top \\ \vdots \\ \tilde{c}_m^\top \end{bmatrix} = AB = \begin{bmatrix} \tilde{a}_1^\top B \\ \vdots \\ \tilde{a}_m^\top B \end{bmatrix}$$

i.e., i th row of C is i th row of A acting (on left) on B

Inner product interpretation

$$c_{ij} = \tilde{a}_i^T b_j = \langle \tilde{a}_i, b_j \rangle$$

i.e., entries of C are inner products of rows of A and columns of B

- ▶ $c_{ij} = 0$ means i th row of A is orthogonal to j th column of B
- ▶ **Gram matrix** of vectors $f_1, \dots, f_n$ defined as $G_{ij} = f_i^T f_j$
(gives inner product of each vector with the others)
- ▶ $G = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix}^T \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix}$