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 ith measurement or sensor reading (which we know)
- x_j is jth parameter to be estimated or determined
- a_{ij} is sensitivity of ith sensor to jth 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 $y = Ax$

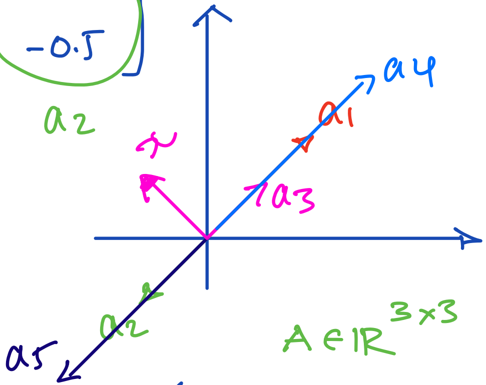
sample problems:

- 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

$$A = \begin{bmatrix} 1 & -0.5 \\ 1 & -0.5 \end{bmatrix}$$

$y = Ax$ $\quad a_1 \quad a_2$



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

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

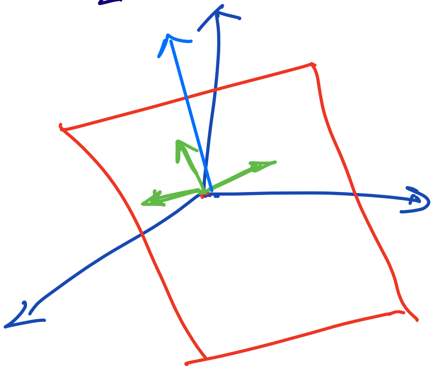
where $a_j \in \mathbb{R}^m$. Then then $y = Ax$ means

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

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

$y = Ax$

$\mathbb{R}^2 \quad \mathbb{R}^5$

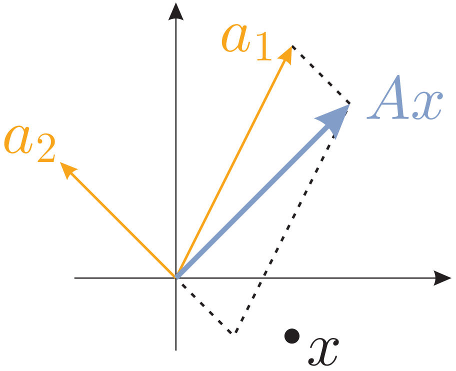


- y is a (linear) combination or mixture of the columns of A
- coefficients of x give coefficients of mixture
- each column of A represents an *actuator*

Geometric interpretation of control

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

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



another example:

$$a_j = Ae_j$$

where e_j is the $j$the unit vector:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

$\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

- jth column of A gives response to unit jth input

Matrix multiplication as inner product with rows

$a^T b$

inner product

write A in terms of its rows:

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

$\tilde{a}_1^T \rightarrow \tilde{a}_1 = \begin{bmatrix} \tilde{a}_{11} \\ \vdots \\ \tilde{a}_{1n} \end{bmatrix}$

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

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

then $y = Ax$ can be written as

$\tilde{a}_{11} x_1 + \tilde{a}_{12} x_2 + \cdots$

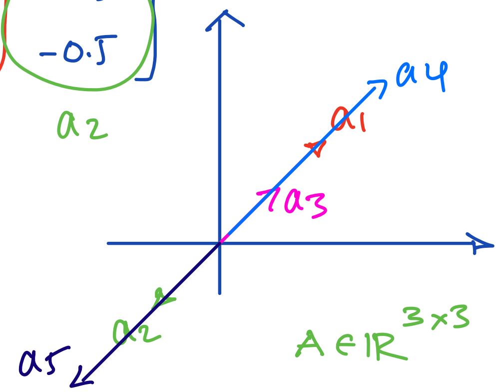
$$y = \begin{bmatrix} \tilde{a}_1^T x \\ \tilde{a}_2^T x \\ \vdots \\ \tilde{a}_m^T x \end{bmatrix}$$

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

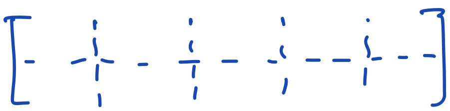
$$x_0 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & -0.5 \\ 1 & -0.5 \end{bmatrix}$$

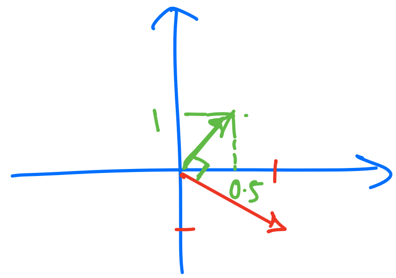
$y = Ax$ $\quad a_1 \quad a_2$



$$\begin{bmatrix} 1 & -0.5 \\ 1 & -0.5 \end{bmatrix} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$A =$ 

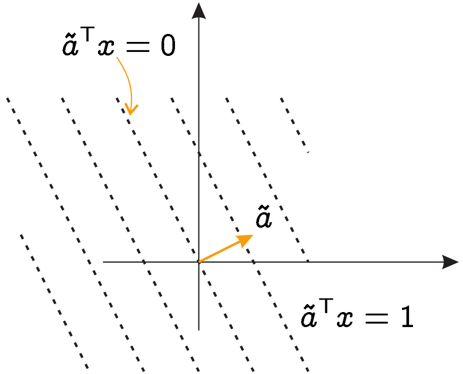
$$\underset{\mathbb{R}^2}{y} = A \underset{\mathbb{R}^5}{x}$$



Geometric interpretation of estimation

$$a_i^\mathsf{T} x = \text{constant}$$

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

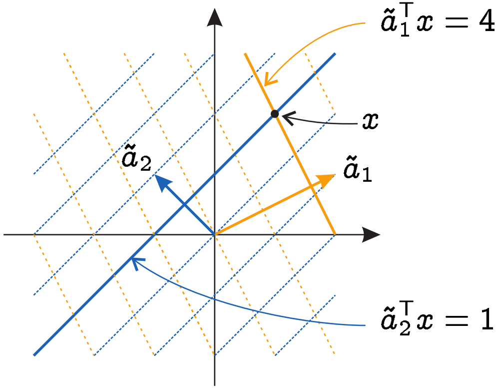


if $Ax = y$ then x is on intersection of hyperplanes $a_i^\mathsf{T} x = y_i$

$a_1^\mathsf{T} x = y_1$

- $A = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$
- $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
- $y = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

$$\begin{cases} 2x_1 + x_2 = 4 \\ -x_1 + x_2 = 1 \end{cases}$$



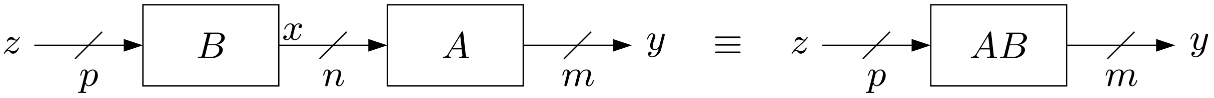
Matrix multiplication as composition

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

$$\begin{bmatrix} I_{2\times 2} & 0_{2\times 3} \\ B & C \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{matrix} \}\, 2 \\ \}\, 3 \end{matrix}$$

$$C = AB, \quad A: m\times n, \quad B: n\times p$$

can write product $C = AB$ as

$$A \begin{bmatrix} b_1 & \cdots & b_p \end{bmatrix}$$

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

$$A \begin{bmatrix} \tilde{b}_1^T \\ \vdots \\ \tilde{b}_n^T \end{bmatrix}$$

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

$$\begin{bmatrix} \tilde{a}_1^T \\ \vdots \\ \tilde{a}_m^T \end{bmatrix} B$$

similarly we can write

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

$$\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} B$$

$$a_1 B \quad \cdots \quad a_n B$$

$$\underbrace{\begin{bmatrix} a_1 \end{bmatrix}}_{m\times 1} \underbrace{\begin{bmatrix} B \end{bmatrix}}_{n\times p}$$

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

Inner product interpretation

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

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

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

$$\underbrace{\begin{bmatrix} \tilde{a}_{i1} & \cdots & \tilde{a}_{in} \end{bmatrix}}_{A} \begin{bmatrix} b_{ij} \\ \vdots \\ b_{nj} \end{bmatrix}$$

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

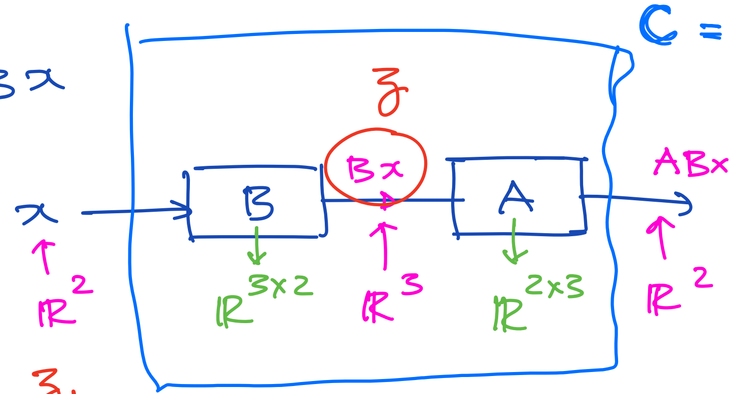
$$A A^{-1} = I = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots & \\ & & & 1 \end{bmatrix}$$

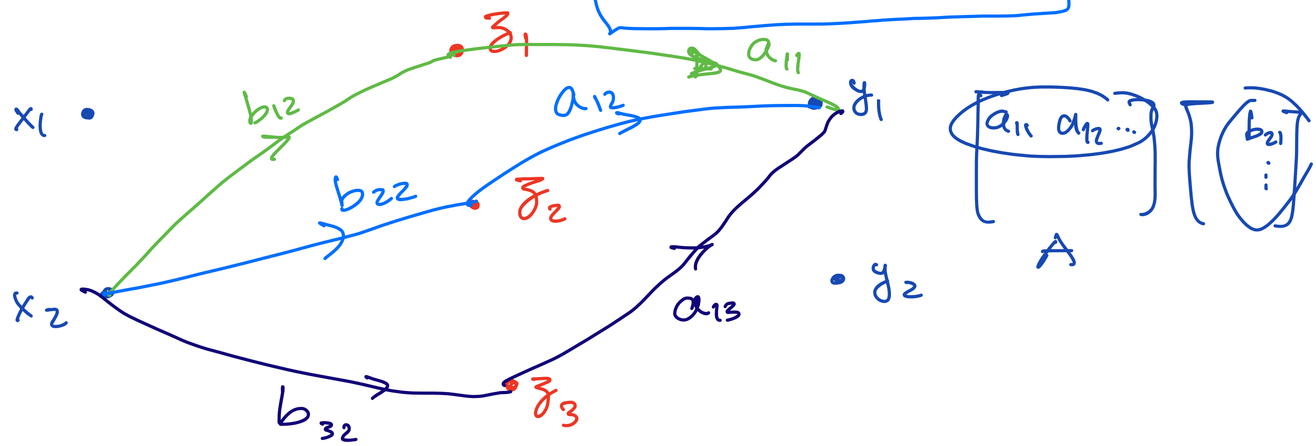
C_{12}

second input → first output

$y = ABx$

$C = AB$





$$a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} = C_{12}$$