2. Linear Algebra Review

- Linear equations
- Application examples
- Control interpretation
- Example: forces on rigid body
- Estimation Interpretation
- Example: navigation
- Block matrices
- Block diagrams
- Range and Null space
- Rank
- Orthogonal matrices

Linear equations

some familiar equations:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_2 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_2 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{aligned}$$

write this as $y = Ax$, where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

this defines a map from $\mathbb{R}^n$ to $\mathbb{R}^m$; this map is *linear*, that is

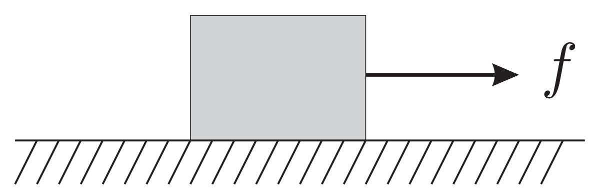
$$A(x+y) = Ax + Ay$$
$$A(\lambda x) = \lambda Ax$$

for any $x, y \in \mathbb{R}^n$ and any $\lambda \in \mathbb{R}$.

Engineering Examples

final position/velocity of mass from applied forces

- unit mass, with zero position/velocity at $t = 0$, subject to force $f(t)$ for $0 \leq t \leq n$
- $f(t) = x_j$ for t in the interval $[j-1, j)$.
  (x is the sequence of applied forces, constant in each interval)
- y_1, y_2 are final position and velocity (i.e. at $t = n$)

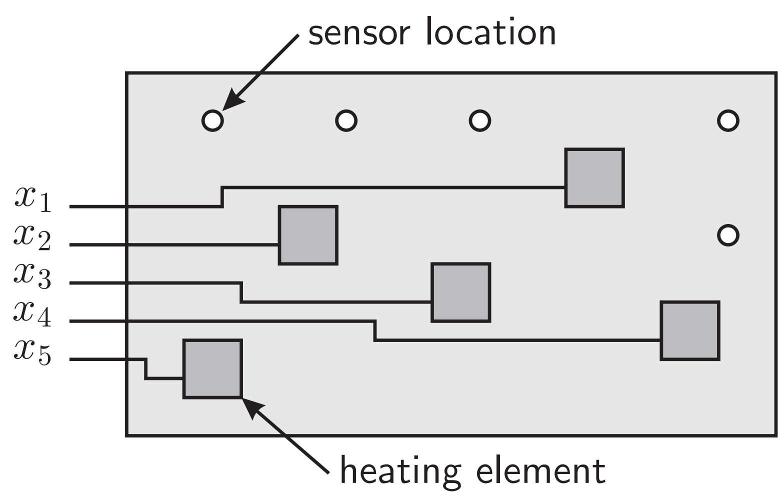


we have $y = Ax$

- a_{1j} gives influence of applied force during $j - 1 \leq t < j$ on final position
- a_{2j} gives influence of applied force during $j - 1 \leq t < j$ on final velocity

heating system with multiple heating elements

- x_j is power of jth heating element
- y_i is change in steady-state temperature at location i
- thermal transport via conduction

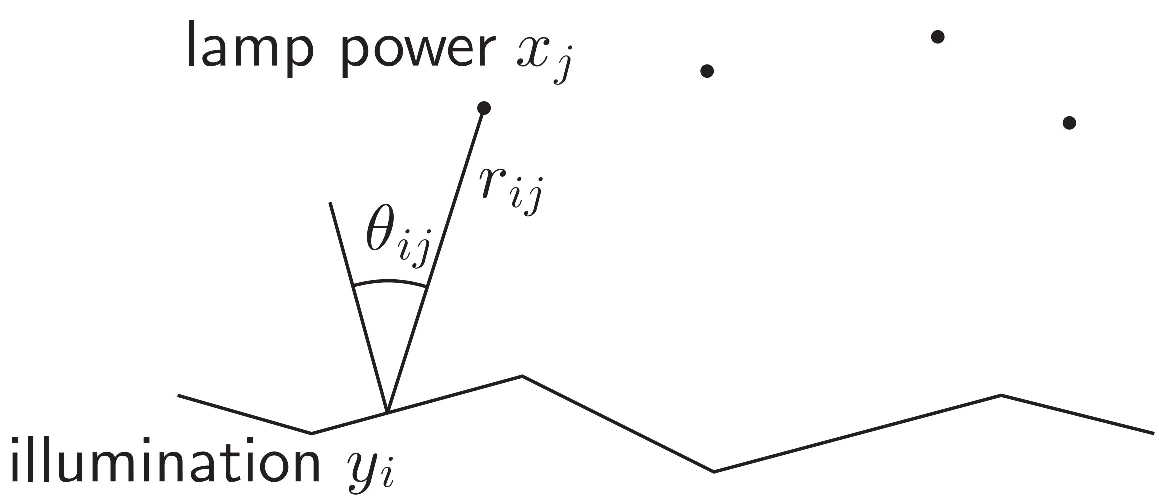


we have $y = Ax$

- a_{ij} gives influence of heater j at location i (in $^\circ/W$)
- jth col. of A gives pattern of steady-state temperature rise due to $1W$ at heater j
- ith row shows how heaters affect location i

illumination with multiple lamps

- n lamps illuminating m (small, flat) patches, no shadows
- x_j is power of jth lamp
- y_i is illumination level of patch i



$y = Ax$, where

- $a_{ij} = r_{ij}^{-2} \cos \theta_{ij}$
- jth column of A shows illumination pattern resulting from lamp j (at $1W$)

signal and interference power in wireless system

n transmitter/receiver pairs

- transmitter j transmits to receiver j (and, inadvertently, to the other receivers)
- p_j is power of jth transmitter
- s_i is receiver signal power of ith receiver
- z_i is receiver interference power of ith receiver
- G_{ij} is path gain from transmitter j to receiver i

we have $s = Ap$ and $z = Bp$ where

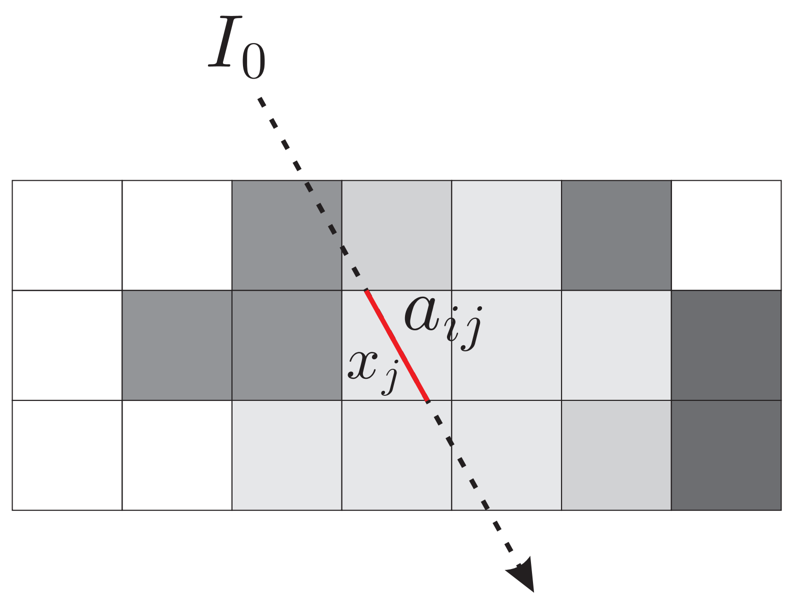
$$a_{ij} = \begin{cases} G_{ii} & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \qquad b_{ij} = \begin{cases} 0 & \text{for } i = j \\ G_{ij} & \text{otherwise} \end{cases}$$

- A is diagonal; B has zero diagonal
- we'd like A 'large', B 'small'

cross-section image reconstruction

tomography or CAT scan, atmospheric/oceanographic remote sensing

- m X-ray beams with intensity I_0
- object is divided into n volume cells ('voxels')
- x_j is density of cell j, $j = 1, \ldots, n$
- $y_i = \log(I_0/I_i)$ where I_i is the measured intensity



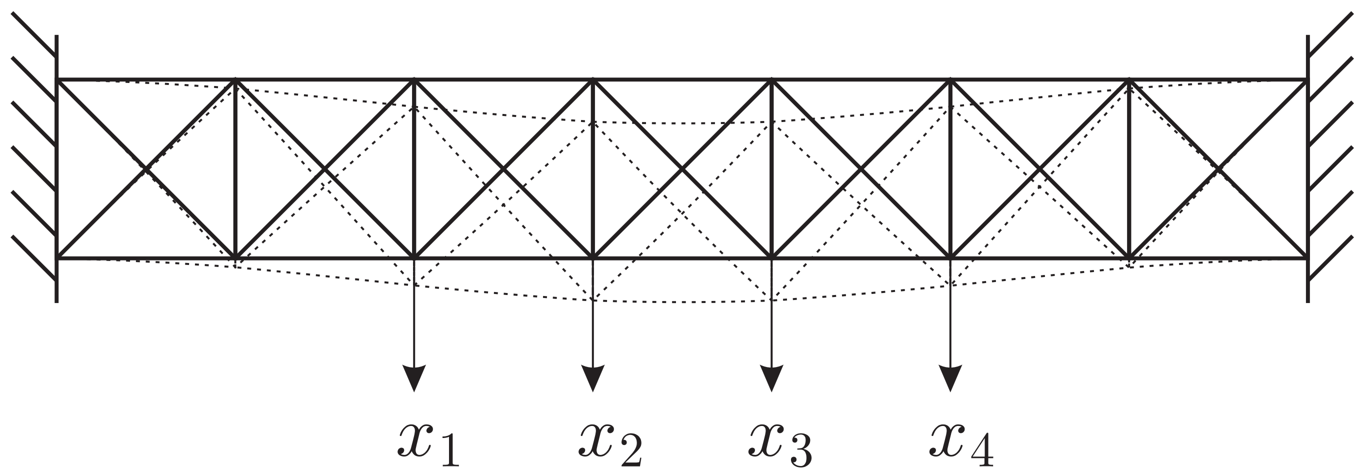
$$I_i = I_0 e^{-\sum_{j=1}^n a_{ij} x_j}$$

$y = Ax$ where

- a_{ij} is length of path of beam j through cell i

linear elastic structure

- x_j is external force applied at some node, in some direction
- y_i is (small) deflection of some node, in some fixed direction



for small displacements, we have $y \approx Ax$

- The matrix A is called the *compliance matrix*
- a_{ij} gives deflection i per unit force j (in m/N)

Control Interpretation of Linear Equations

we have the equation

$$y = Ax$$

- x is a vector of inputs or design parameters we choose
- y is the vector of results or outcomes
- a_{ij} is the sensitivity of the ith outcome to the jth parameter

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 the specifications)
- among all x's that satisfy $y = y_{\text{des}}$, find a small one
  (i.e., find a small or efficient x that meets specifications)

control interpretation via columns

write A in terms of its columns

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

where $a_i \in \mathbb{R}^m$ for all i. Then $y = Ax$ means

$$\begin{aligned} y &= x_1 a_1 + x_2 a_2 + \cdots + x_n a_n \\ &= \sum_{j=1}^{n} x_j a_j \end{aligned}$$

here each a_j is a vector

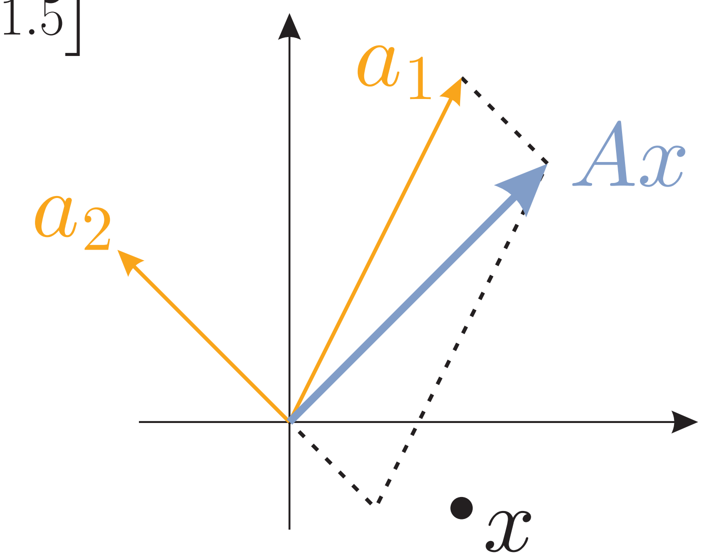
- usually we think of the matrix A as acting on x to produce y
- for control, it makes sense to think of x as acting on A to produce y

each column of A represents an *actuator*

geometric interpretation of control

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

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



another example:

$$a_j = Ae_j$$

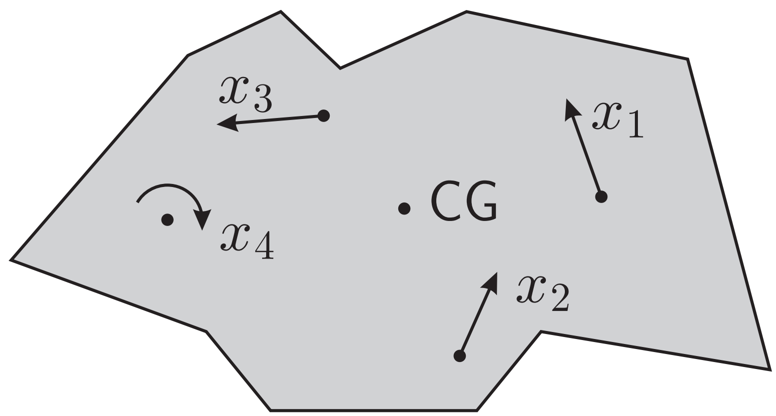
where e_j is the $j$the unit vector:

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

- jth column of A gives response to unit jth input

application example: total force/torque on rigid body

- x_j is external force/torque applied at some point/direction/axis
- y is resulting total force and torque on body. six real numbers:
  y_1, y_2, y_3 are x, y, z components of total force
  y_4, y_5, y_6 are x, y, z components of total torque



we have $y = Ax$

- A depends on geometry (of applied forces and torques with respect to center of gravity CG)
- jth column gives resulting force and torque for unit force/torque j

Estimation Interpretation of Linear Equations

we also use linear equations to describe *estimation problems*; again we have the equation

$$y = Ax$$

- y_i is the ith measurement or sensor reading
- x_j is the jth parameter to be estimated or determined
- a_{ij} is the sensitivity of the ith sensor to the jth parameter

sample problems

- given y_{meas}, find x
- find all x that result in y_{meas}
  (i.e., all x *consistent* with measurements)
- if there is no x such that $y_{\text{meas}} = Ax$, find x such that $y_{\text{meas}} \approx Ax$
  (i.e., if the sensor readings are inconsistent, find x which is almost consistent)

estimation interpretation via rows

write A in terms of its rows

$$A = \begin{bmatrix} b_1^T \\ b_2^T \\ \vdots \\ b_m^T \end{bmatrix}$$

then

$$y = \begin{bmatrix} b_1^T x \\ b_2^T x \\ \vdots \\ b_m^T x \end{bmatrix}$$

where $b_i \in \mathbb{R}^n$, so that

- y_i is the scalar product of b_i with x
- if b_i is a unit vector, then y_i is the *component* of x in the direction b_i
- think of A as acting on x to produce y

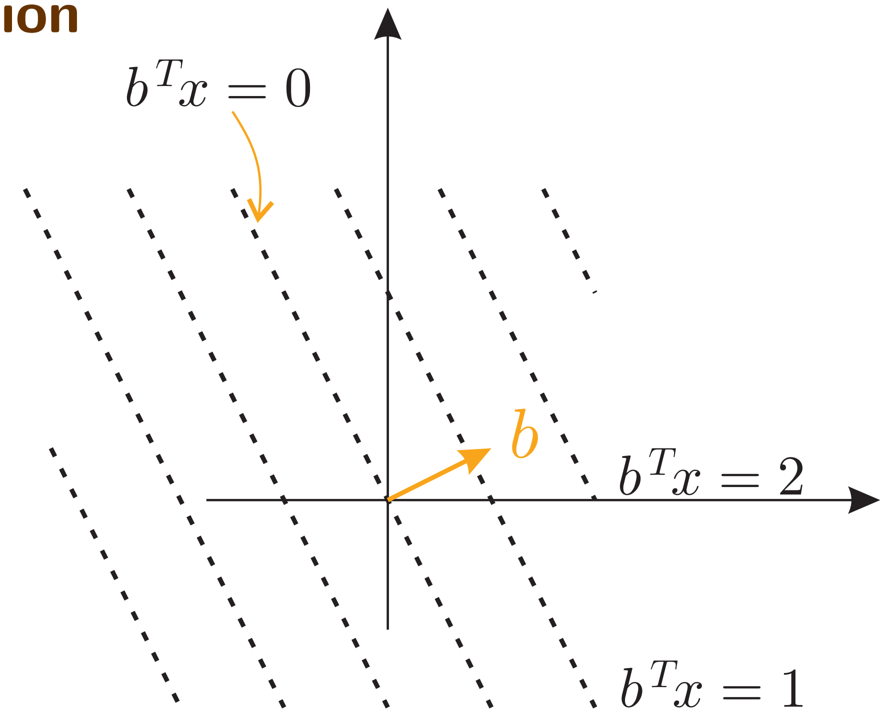
in particular,

each row of A represents a *sensor*

geometric interpretation of estimation

$$b_i^T x = \text{constant}$$

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

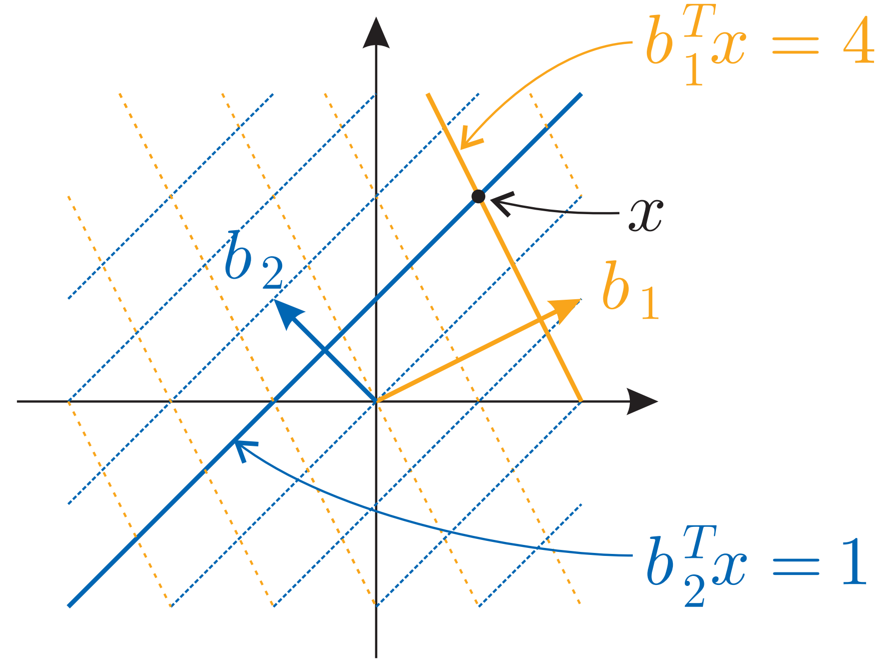


if $Ax = y$ then x is on intersection of hyperplanes $b_i^T x = y_i$
example:

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

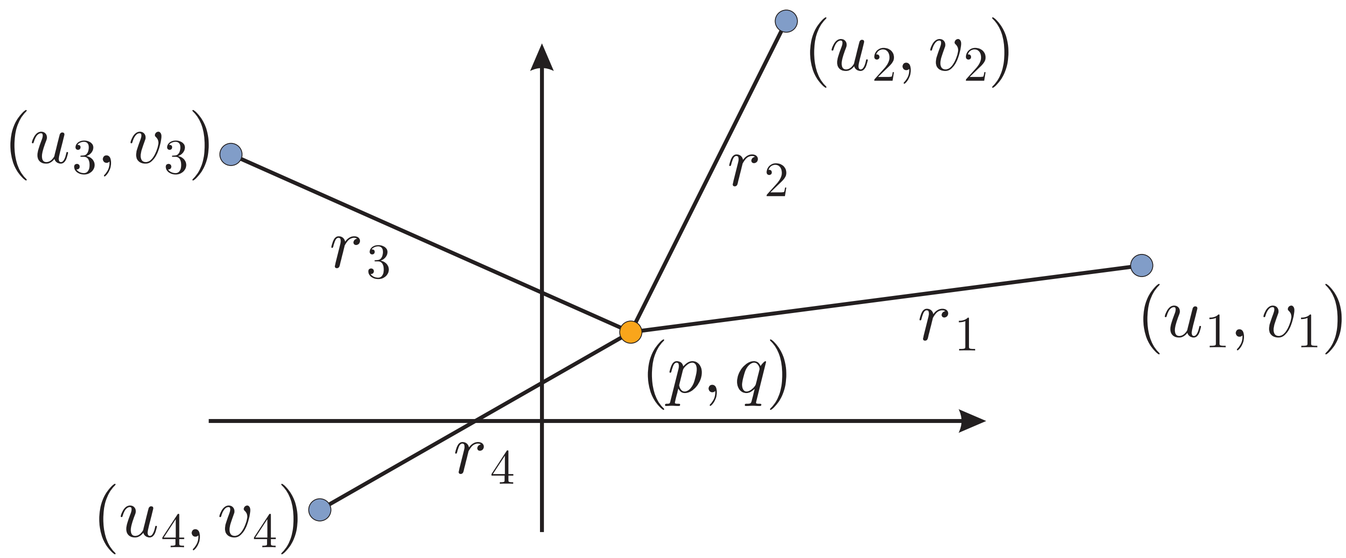
$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

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



example: navigation

$(p, q) \in \mathbb{R}^2$ is our location, and we measure distances r_i to m beacons at points (u_i, v_i)



$$r_i = \sqrt{(p - u_i)^2 + (q - v_i)^2} = f_i(p, q)$$

navigation example continued

Taylor expansion is

$$\begin{aligned} r_i &= f_i(0,0) + \begin{bmatrix} \frac{\partial f_i(0,0)}{\partial p} \\ \frac{\partial f_i(0,0)}{\partial q} \end{bmatrix}^T \begin{bmatrix} p \\ q \end{bmatrix} \\ &= \sqrt{u_i^2 + v_i^2} - \frac{1}{\sqrt{u_i^2 + v_i^2}} \begin{bmatrix} u_i \\ v_i \end{bmatrix}^T \begin{bmatrix} p \\ q \end{bmatrix} \end{aligned}$$

assume p, q are small compared to u_i, v_i. then $y \approx Ax$ where

- $A \in \mathbb{R}^{m \times 2}$, ith row of A is the transpose of unit vector in the direction of beacon i
- $y = \begin{bmatrix} \sqrt{u_1^2 + v_1^2} - r_1 \\ \vdots \\ \sqrt{u_m^2 + v_m^2} - r_m \end{bmatrix}$ measured vector of distances
- $x = \begin{bmatrix} p \\ q \end{bmatrix}$ our location

Block Matrices and Vectors

if P, Q, R, S are

$$P = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \quad Q = \begin{bmatrix} 7 \\ 1 \end{bmatrix} \quad R = \begin{bmatrix} 3 & 3 \end{bmatrix} \quad S = 7$$

then

$$A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \qquad \text{means} \qquad A = \begin{bmatrix} 1 & 2 & 7 \\ 3 & 1 & 1 \\ 3 & 3 & 7 \end{bmatrix}$$

A is called a *partitioned matrix* or a *block matrix*.

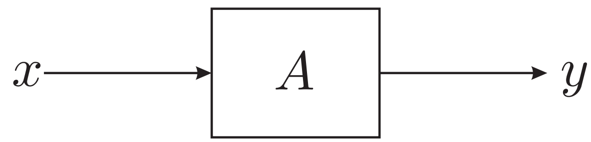
For this to make sense, we need the dimensions of P, Q, R, S to be *compatible*

for vectors:

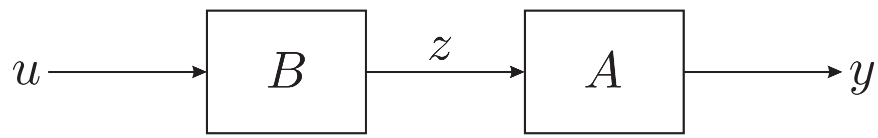
$$a = \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} \qquad \Longrightarrow \qquad \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 3 \\ 6 \\ 2 \end{bmatrix}$$

Block Diagrams

we can also represent $y = Ax$ by the block diagram:



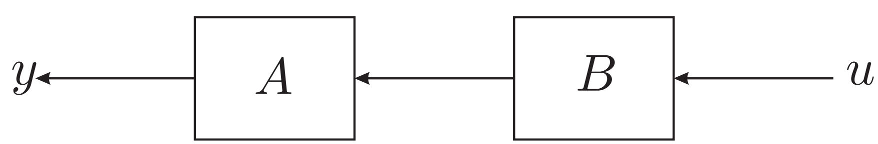
if $y = ABu$ then



so $z = Bu$, $y = Az$ hence $y = ABu$.

Note the order of the blocks. In general $AB \neq BA$.

so we often draw block diagrams right-to-left:



block diagrams and block matrices

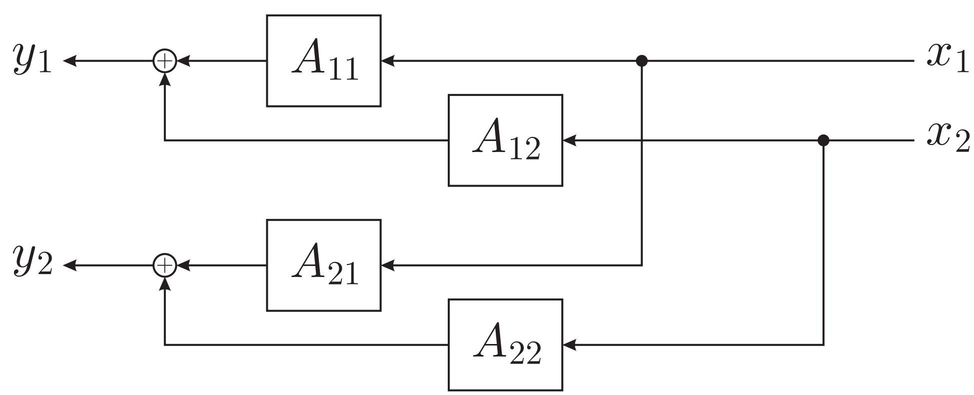
suppose

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

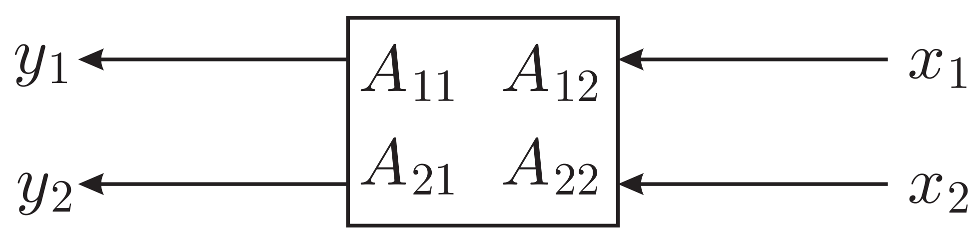
where $A_{11} \in \mathbb{R}^{m_1 \times n_1}$, $A_{12} \in \mathbb{R}^{m_1 \times n_2}$, $A_{21} \in \mathbb{R}^{m_2 \times n_1}$, $A_{22} \in \mathbb{R}^{m_2 \times n_2}$. If

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

we draw this as:



convenient shorthand:



perp notation

- for vectors x, y,

$$x \perp y \qquad \text{means} \qquad x^T y = 0$$

- if S is a subspace

$$x \perp S \qquad \text{means} \qquad x^T y = 0 \text{ for all } y \in S$$

- if S and T are both subspaces

$$T \perp S \qquad \text{means} \qquad x^T y = 0 \text{ for all } x \in T, y \in S$$

- if S is a subspace, the *orthogonal complement* of S

$$S^\perp = \{\, x \mid x \perp S \,\}$$

  is the set of all vectors perpendicular to S.

- one can show $S^{\perp\perp} = S$

The Range

$$\mathrm{range}(A) = \{\, Ax \mid x \in \mathbb{R}^n \,\}$$

for *control* problems:

$$\begin{aligned}\mathrm{range}(A) &= \text{set of possible outputs of } y = Ax \\ &= \text{span of columns of } A\end{aligned}$$

the range is also called the *column space* or the *image* of A

the range is important in control problems because

The equation $y = Ax$ has a solution x	$\iff$	$y \in \mathrm{range}(A)$

If we want a solution for all $y \in \mathbb{R}^m$, then we need $\mathrm{range}(A) = \mathbb{R}^m$. Equivalently

- the columns of A span $\mathbb{R}^m$.
- the rows of A are linearly independent.

The Null Space

$$\mathrm{null}(A) = \{\, x \in \mathbb{R}^n \mid Ax = 0 \,\}$$

for *estimation* problems:

$$\begin{aligned}\mathrm{null}(A) &= \text{set of unknowns which produce zero sensor output}\\ &= \text{set of vectors orthogonal to } \textit{all} \text{ rows of } A\end{aligned}$$

the null space is also called the *kernel* of A.

the null space is important in estimation problems because

> if x_0 is one solution to $y = Ax$, then the set of all solutions is
> $$\{\, x_0 + z \mid z \in \mathrm{null}(A) \,\}$$

i.e., we can *unambiguously* determine x from y if and only if $\mathrm{null}(A) = \{0\}$. Equivalently

- the columns of A are linearly independent.
- the rows of A span $\mathbb{R}^n$.
- $\mathrm{range}(A^T) = \mathbb{R}^n$.

Rank

the *rank* of a matrix is the dimension of its column space. i.e.,

$$\operatorname{rank}(A) = \dim \operatorname{range}(A)$$

Notes

- the dimension of the column space equals the dimension of the row space. i.e.,

$$\operatorname{rank}(A) = \operatorname{rank}(A^T)$$

- an $m \times n$ matrix A is called *full rank* if

$$\operatorname{rank}(A) = \min\{m, n\}$$

- A is called *skinny* if $n < m$ and *fat* if $n > m$
- A is called *full column rank* if its columns are linearly independent

Properties of Rank

an important property of rank is

$$\boxed{\dim \operatorname{range}(A) + \dim \operatorname{null}(A) = n}$$

we interpret this as *conservation of dimension*. Of n input dimensions, every one is either mapped to zero or mapped to the output.

Therefore if $\operatorname{range} A = \mathbb{R}^m$ we must have $n \geq m$. And if $\operatorname{null} A = \{0\}$ we must have $n \leq m$.

Therefore

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$$

Orthogonal Matrices

a square matrix $A \in \mathbb{R}^{n \times n}$ is called *orthogonal* if

$$A^T A = I$$

Properties

- if $y = Ax$ then

$$\begin{aligned} \|y\|^2 &= \|Ax\|^2 \\ &= (Ax)^T Ax \\ &= x^T A^T Ax \\ &= \|x\|^2 \end{aligned}$$

  therefore A preserves the lengths of vectors. It follows that

$$\|Ax_1 - Ax_2\| = \|x_1 - x_2\|$$

  and so A *preserves distances* between vectors. A is called an *isometry*.

- A is invertible, since $A^{-1} = A^T$. Therefore $AA^T = I$ also.

Orthonormal Bases

if $q_1, q_2, \ldots, q_n \in \mathbb{R}^n$ are orthonormal, then the matrix $U = \begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix}$ is orthogonal.

to show this, let $W = U^T U$.

the i, j element of W is

$$\begin{aligned} W_{ij} &= q_i^T q_j \\ &= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

therefore $W = I$.