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
- Orthogonal matrices
- Rank

2. Linear Algebra Review
Linear equations

Some familiar equations:

\[ y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \]

\[ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \]

\[ \vdots \]

\[ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \]

Write this as:

\[ y = Ax \]

where

\[ A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \]

\[ x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \]

\[ y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \]

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 \( \lambda \in \mathbb{R} \).

Some familiar equations:

\[ u_1x_1 + \cdots + u_nx_n = \, \]

\[ \vdots \]

\[ u_1x_1 + \cdots + u_nx_n = \, \]

\[ \vdots \]

\[ u_1x_1 + \cdots + u_nx_n = \, \]

\[ 1 \]
\( f \) gives influence of applied force during \( \ell > t \) on final position.

\( \forall \ell \) gives influence of applied force during \( \ell \) on final position.

\( y_1, y_2 \) are final position and velocity (i.e. at \( t = n \))

\( (u = t) \) are final position and velocity (i.e. at \( t = n \))

\( x \) is the sequence of applied forces, constant in each interval.

\( \forall \ell \), for \( \ell \) in the interval \( \ell \), \( x = (t)f \)

\( u \geq t \geq 0 \) for \( (t)f \) force, subject to force from applied forces

Final position/velocity of mass from applied forces

**Examples**
heating system with multiple heating elements

• \( x_j \) is power of \( j \)th heating element
• \( y_i \) is change in steady-state temperature at location \( i \)

\[ M / \text{sensor} \rightarrow \text{heating element} \]

\[ y_i = \text{change in steady-state temperature at location } i \]
\[ x_j = \text{power of } j \text{th heating element} \]

Thermal transport via conduction

We have

\[ Ax = y \]
Multiple lamps illuminating small, flat patches with no shadows:

- $n$ lamps illuminating $m$ patches.
- $x_j$ is the power of the $j$th lamp.
- $y_i$ is the illumination level of patch $i$.

The relationship between the $y_i$ values and the $x_j$ powers can be described by the equation:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

where $A$ is an $m \times n$ matrix, $A = [a_{ij}]$, and $y_i$ is the vector of illumination levels for each patch.

Each column of $A$ represents the illumination pattern resulting from the respective lamp (at 1W).
transmitter \( j \) transmits to receiver \( i \) (and, inadvertently, to the other receivers)

\( p_j \) is power of \( j \)th transmitter

\( s_i \) is receiver signal power of \( i \)th receiver

\( z_i \) is receiver interference power of \( i \)th receiver

\( G_{ij} \) is path gain from transmitter \( j \) to receiver \( i \)

\( d BG = z \) and \( d A V = s \) where

\[
\begin{pmatrix}
G_{ii} & 0 \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
p_j \\
0
\end{pmatrix}
= \begin{pmatrix}
p_j \\
0
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & G_{ij} \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
p_j \\
0
\end{pmatrix}
= \begin{pmatrix}
p_j \\
0
\end{pmatrix}
\]

\( A \) is diagonal; \( B \) has zero diagonal

we'd like \( A \), large; \( B \), small

\( d BG = z \) and \( d A V = s \) where

\( G_{ii} \) is path gain from transmitter \( j \) to receiver \( i \)

\( z_i \) is receiver interference power of \( i \)th receiver

\( s_i \) is receiver signal power of \( i \)th receiver

\( p_j \) is power of \( j \)th transmitter

\( d_p \) is power of \( j \)th transmitter

\( d \) transmitter/receiver pairs

\( s \) transmitter/receiver pairs in wireless system
cross-section image reconstruction

- Object is divided into $n$ volume cells (voxels).
- $x \in A$ is density of cell $x$, where $x \in A = \{x\}$.
- $y_i = \log(I_0/I_i)$ where $I_i$ is the measured intensity.
- $A$ is the set of all voxels.
- $a_{ij}$ is length of path of beam $j$ through cell $i$.
- $\mathcal{D}$ is the set of all paths.

\[ y = Ax \]

\[ \int_{\mathcal{D}} \frac{I_0}{e^{a_{ij}x_j}} I = y \]
\( \begin{align*}
\vec{x} & \approx \vec{f} \\
A & \text{ is called the compliance matrix}
\end{align*} \)

The matrix \( A \) is called the compliance matrix.

For small displacements, we have \( Ax \approx f \).

\begin{align*}
x_1 & \quad x_2 \\
x_3 & \quad x_4
\end{align*}
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 \( i \)th outcome to the \( j \)th 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)

We choose \( x \) as a vector of inputs or design parameters. We choose \( \mathbf{f} \) as the vector of results or outcomes. \( a_{ij} \) is the sensitivity of the \( j \)th outcome to the \( i \)th parameter.
For control, it makes sense to think of $x$ as acting on $A$ to produce $y$. Usually we think of the matrix $A$ as acting on $x$ to produce $y$.

Here each $a_j$ is a vector:

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

Then $x \mathcal{A} = y$ means

$$\begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix} = A$$

Each column of $A$ represents an actuator.
another example:

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix}
\end{bmatrix} + \begin{bmatrix}
1 \\
1
\end{bmatrix} = \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

Geometric interpretation of control

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix} = h \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} = x \\
\begin{bmatrix}
1 & 2 \\
1 & -1
\end{bmatrix} = A
\]

\[\text{where } e_j \text{ is the } j\text{th unit vector:}
\]

\[
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}, \begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix}, \ldots, \begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix}
\]

\[
\text{another example:}
\]

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix} \begin{bmatrix}
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix}
\end{bmatrix} + \begin{bmatrix}
1 \\
1
\end{bmatrix} = \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix} = h \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} = x \\
\begin{bmatrix}
1 & 2 \\
1 & -1
\end{bmatrix} = A
\]

Geometric interpretation of control

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix} = h \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} = x \\
\begin{bmatrix}
1 & 2 \\
1 & -1
\end{bmatrix} = A
\]

Geometric interpretation of control

\[
\begin{bmatrix}
0.5 & 1 \\
0.5 & 1
\end{bmatrix} = h \\
\begin{bmatrix}
0.5 \\
0.5
\end{bmatrix} = x \\
\begin{bmatrix}
1 & 2 \\
1 & -1
\end{bmatrix} = A
\]

Geometric interpretation of control
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: $y_1, y_2, y_3, y_4, y_5, y_6$ are $x, y, z$ components of total force.
- $y_1, y_2, y_3$ are $x, y, z$ components of total torque.
- $y$ is resulting total force and torque on body.

we have

$Ax = y$ where $A$ depends on geometry (of applied forces and torques with respect to center of gravity $CG$)

$A$ is $j$th 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

\[ y = Ax \]

• \( y_{\text{meas}} \) is the \( i \)th measurement or sensor reading
• \( x_j \) is the \( j \)th parameter to be estimated or determined
• \( a_{ij} \) is the sensitivity of the \( i \)th sensor to the \( j \)th parameter

sample problems

\[ x \text{ consistent with measurements} \]

(i.e., all \( x \) that result in \( y_{\text{meas}} \))

\[ y_{\text{meas}} \approx Ax \]

(i.e., if the sensor readings are inconsistent, find \( x \) which is almost consistent)

\[ x \text{ consistent with measurements} \]

(i.e., all \( x \) that result in \( y_{\text{meas}} \))

Given \( y_{\text{meas}} \), find \( x \)

we also use linear equations to describe estimation problems; again we have the equation
each row of $A$ represents a sensor

In particular,

- think of $A$ as acting on $x$ to produce $\hat{y}$
- if $y_i$ is a unit vector, then $y_i^T$ is the component of $x$ in the direction $q_i$
- $y_i$ is the scalar product of $q_i$ with $x$

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

$$\begin{bmatrix} x_{LQ}^u \\ x_{LQ}^v \\ \vdots \\ x_{LQ}^z \\ x_{LQ}^w \\ \vdots \\ x_{LQ}^z \\ x_{LQ}^w \end{bmatrix} = \hat{y}$$

then

$$\begin{bmatrix} w_{LQ}^u \\ \vdots \\ z_{LQ}^v \\ \vdots \\ z_{LQ}^z \\ \vdots \\ w_{LQ}^w \end{bmatrix} = A$$

Write $A$ in terms of its rows estimation interpretation via rows
$I = x_L^T q$

$\forall = x_L^T q$

$\exists = x_L^T q$

$0 = x_L^T q$

$\mathbf{Ax} = \mathbf{y}$

If $\mathbf{x}$ is on intersection of hyperplanes $\mathbf{b}_i x = \mathbf{y}_i$.

Example:

If $\mathbf{y} = x_L^T q$ then $\mathbf{h} = x_L^T \mathbf{A}$

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

Geometric Interpretation of Estimation

S. Lall, Stanford 2007.09.27.02
\[(b',d)'f = \frac{1}{\ell}(\nu - b) + \frac{1}{\ell}(\nu - d) \wedge = \nu\]

In navigation, our location is \((p, q) \in \mathbb{R}^2\) and we measure distances \(r_i\) to \(m\) beacons at points \((u_i, v_i)\). For example, navigation is our location, and we measure distances to \(m\) beacons at points \((\nu, \nu)\).
Taylor expansion is
\[
\begin{bmatrix}
\bar{b} \\
\bar{d}
\end{bmatrix} = \begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix} \approx \begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix} = \begin{bmatrix}
\frac{\bar{b}_Q}{(0,0)^T f} \\
\frac{\bar{d}_Q}{(0,0)^T f}
\end{bmatrix} + (0,0)^T f = \bar{u}
\]

Measured vector of distances
\[
\begin{bmatrix}
\frac{\vec{u} \cdot \bar{a} + \vec{u} \cdot \bar{n}}{2} \\
\frac{\vec{v} \cdot \bar{a} + \vec{v} \cdot \bar{n}}{2} \\
\frac{\vec{v} \cdot \bar{a} + \vec{v} \cdot \bar{n}}{2}
\end{bmatrix}
\]

where \(\bar{u}, \bar{v}, \bar{a}, \bar{n}\) are small compared to \(\vec{u}, \vec{v}, \vec{a}, \vec{n}\). Then, our location
\[
\begin{bmatrix}
\bar{b} \\
\bar{d}
\end{bmatrix} = \begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix}
\]

Assume \(p, q\) are small compared to \(u_i, v_i\). Then
\[
\begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix} \approx \begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix}
\]

Our location
\[
\begin{bmatrix}
\bar{b} \\
\bar{d}
\end{bmatrix} = \begin{bmatrix}
\vec{u} \\
\vec{v}
\end{bmatrix}
\]

is the transpose of unit vector in the direction of beacon \(i\).
Block Matrices and Vectors

If \( P, Q, R, S \) are

\[
\begin{bmatrix}
1 & 2 & 3 \\
6 & 7 & 3 \\
1 & 2 & 7
\end{bmatrix} = \begin{bmatrix} q \\ p \end{bmatrix} \quad \iff \quad \begin{bmatrix}
1 & 2 \\
6 & 3 \\
1 & 7
\end{bmatrix} = q \quad , \quad \begin{bmatrix} 1 \\ 3 \end{bmatrix} = p
\]

then

\[
\begin{bmatrix}
7 & 3 & 3 \\
1 & 1 & 3 \\
7 & 2 & 1
\end{bmatrix} = \begin{bmatrix} \mathcal{A} \\ \mathcal{B} \end{bmatrix} \quad \iff \quad \begin{bmatrix} \mathcal{S} & \mathcal{H} \\ \mathcal{O} & \mathcal{D} \end{bmatrix} = \begin{bmatrix} \mathcal{A} \\ \mathcal{B} \end{bmatrix}
\]

A is called a partitioned matrix or a block matrix.

For vectors, we need the dimensions of \( P, Q, R, S \) to be compatible:

\[
\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \quad , \quad \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix}
\]

Block Matrices and Vectors

If \( P, Q, R, S \) are
we can also represent \( y = Ax \) by the block diagram:

\[
\begin{array}{c}
\text{n} \\
\bullet \quad B \\
\text{V} \\
\bullet \quad \text{y}
\end{array}
\]

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

\[
\begin{array}{c}
\text{n} \\
\bullet \quad B \\
\text{V} \\
\bullet \quad \text{y}
\end{array}
\]

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

\[
\begin{array}{c}
\text{B} \\
\bullet \quad A \\
\text{z} \\
\bullet \quad \text{B}
\end{array}
\]

so \( A B n = y \), hence \( A z = y \), \( B u = z \).

If \( B u = z \) then \( A B n = y \).

\[
\begin{array}{c}
\text{B} \\
\bullet \quad A \\
\text{x} \\
\bullet \quad \text{y}
\end{array}
\]

we can also represent \( y = x \) by the block diagram:

\[
\begin{array}{c}
\text{B} \\
\bullet \quad A \\
\text{x} \\
\bullet \quad \text{y}
\end{array}
\]

Block Diagrams
block diagrams and block matrices

Suppose

\[
\begin{bmatrix}
    \mathbf{y}_1 \\
    \mathbf{y}_2
\end{bmatrix} = \begin{bmatrix}
    \mathbf{A}_{11} & \mathbf{A}_{12} \\
    \mathbf{A}_{21} & \mathbf{A}_{22}
\end{bmatrix} \begin{bmatrix}
    \mathbf{x}_1 \\
    \mathbf{x}_2
\end{bmatrix}
\]

where \( \mathbf{A}_{i,j} \in \mathbb{R}^{m \times n} \), for all \( i, j \).
perp notation

S = ⊥⊥S

is the set of all vectors perpendicular to S.

\{ S \perp x \mid x \}\ = \ ⊥S

if S is a subspace, the orthogonal complement of S

S \subseteq \mathcal{H} \forall x \mid \forall \ L \subseteq \mathcal{H} \exists x \mid 0 = h_L x \means S \perp L

if S and L are both subspaces

S \subseteq \mathcal{H} \forall x \mid 0 = h_L x \means S \perp x

if S is a subspace

0 = h_L x \means \mathcal{H} \perp x

for vectors x, \ \mathcal{H} \perp x
The Range

\[ \text{range}\ (A) = \{ Ax \mid x \in \mathbb{R}^n \} \]

For control problems:

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

Equivalently:

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

\[ (\forall y \in \text{range}(A)) \iff (\forall x \in \mathbb{R}^n) \exists Ax = y \]

The equation \( y = Ax \) has a solution \( x \) if and only if \( y \) is in the range of \( A \).

The range is important in control problems because:

- the range is also called the column space or the image of \( A \) = span of columns of \( A \) = \( \{ u \in \mathbb{R}^n \mid uA \} = \text{range}(A) \)

For control problems:

\[ \{ u \in \mathbb{R}^n \mid uA \} = \text{range}(A) \]

The Range
The Null Space

\[ \text{null}(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \} \]

for estimation problems:

- \text{null}(A) = \text{set of unknowns which produce zero sensor output}
- \text{null}(A) = \text{set of vectors orthogonal to all rows of } A

The null space is also called the kernel of \( A \).

\[ \{ (\forall)_{\text{null}} \ni z \mid z + 0x \} \]

If \( x \) is one solution to \( y = Ax \), then the set of all solutions is

\[ \{ 0 = xA \mid u \in \mathbb{R} \} = (A)_{\text{null}} \]

The Null Space
The rank of a matrix is the dimension of its column space, i.e.,

\[ \text{rank}(A) = \dim \text{range}(A) \]

Notes

- A is called full column rank if its columns are linearly independent.
- A is called skinny if \( m > n \) and fat if \( m < n \).
- A is called full rank if \( \text{rank}(A) = \min\{m,n\} \).
- 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

\[
\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}
\]

Therefore

\[
u \leq m.
\]

Therefore if \( \text{range } A = \mathbb{R}^m \) we must have \( u \leq m \). And if \( \text{null } A = \{0\} \) we must have \( u \geq m \).

We interpret this as \textit{conservation of dimension}. Of \( n \), input dimensions, every one is either mapped to zero or mapped to the output. Of \( m \), output dimensions, at most \( m \) are nonzero.

\[
u = (\forall) \dim \text{range } A + (\forall) \dim \text{null } A
\]

An important property of rank is
Orthogonal Matrices

A square matrix \( A \in \mathbb{R}^{n \times n} \) is called orthogonal if

\[ A^T A = I. \]

Also, \( A \) is invertible, since \( A^{-1} = A^T \). Therefore \( A \) preserves the lengths of vectors. \( A \) is called an isometry.

\[ \|Ax - x\| = \|x - x\| \]

Therefore \( A \) preserves distances between vectors. It follows that

\[ \|Ax\| = \|x\| \]

\[ xV_L V_L x = \]

\[ xV_L (xV) = \]

\[ \|xV\| = \|x\| \]

If \( f \) then

Properties

\[ I = V_L V \]

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

Orthogonal Matrices
Orthonormal Bases

If \( q_1, q_2, \ldots, q_n \in \mathbb{R}^n \) are orthonormal, then the matrix

\[
U = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}
\]

is orthogonal.

To show this, let

\[
W = U^T U.
\]

The \( ij \) element of \( W \) is

\[
W_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}
\]

Therefore

\[
I = W
\]