Linear Algebra Review

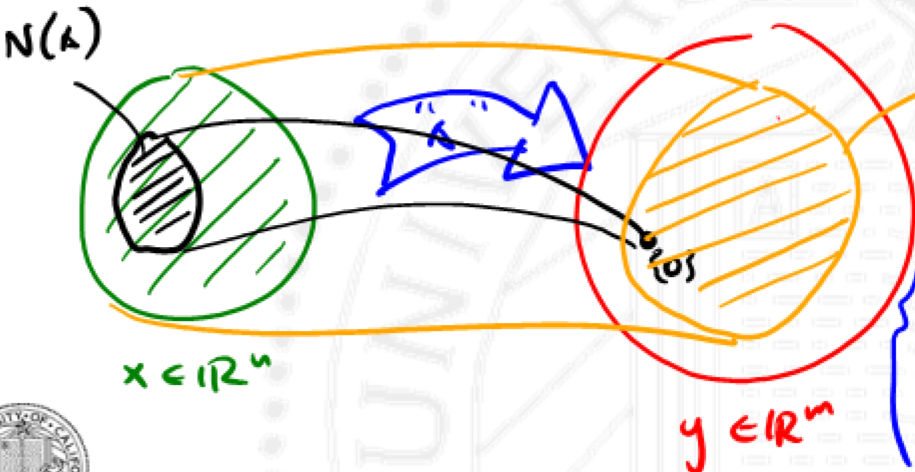
Gabriel Hugh Elkaim

Winter 2016

Range of a Matrix

$\mathcal{R}(A)$: Range of A

$$\mathcal{R}(A) \triangleq (Ax \mid x \in \mathbb{R}^n) \in \mathbb{R}^m$$



$\mathcal{R}(A)$

- Set of vectors that can be "hit" by $y = Ax$
- span$\{a_1 \dots a_m\}$
- set of vectors $\{y\}$ for which $y = Ax$ can be solved

"Onto" Matrices

A is called "onto" if $R(A) = \mathbb{R}^m$

- $y = Ax$ can be solved for <u>any</u> y.
- Columns of A $\{a_1, \dots, a_n\}$ span $\mathbb{R}^m$
- A has a <u>RIGHT</u> inverse $B \in \mathbb{R}^{n \times m}$ $AB = I$
- rows of A are independent $\{\tilde{a}_i\}$ independent

$N(A^T) = \{0\}$

$\det(AA^T) \neq 0$

$y = Ax \quad x = B y_{des}$

$y = \underbrace{AB}_{I} y_{des} \rightarrow y = y_{des}$

Interpretations of Range

$Ax = \sum_{i=1}^{n} x_i a_i = \text{span}\{a_1 \cdots a_n\}$ $\quad A \in \mathbb{R}^{m \times n}$

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

$v \in R(A) \quad w \notin R(A)$

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

image of $\mathbb{R}^n$

$y = Ax$ ← measurement

$R(A)$

$y = v$ possible or consistent sensor/measurement

$y = w$ impossible sensor measurement / inconsistent signal

$y = Ax$ is an output/result

$R(A)$

$y = v$ is a possible result or output

$y = w$ is an impossible or inadmissible result

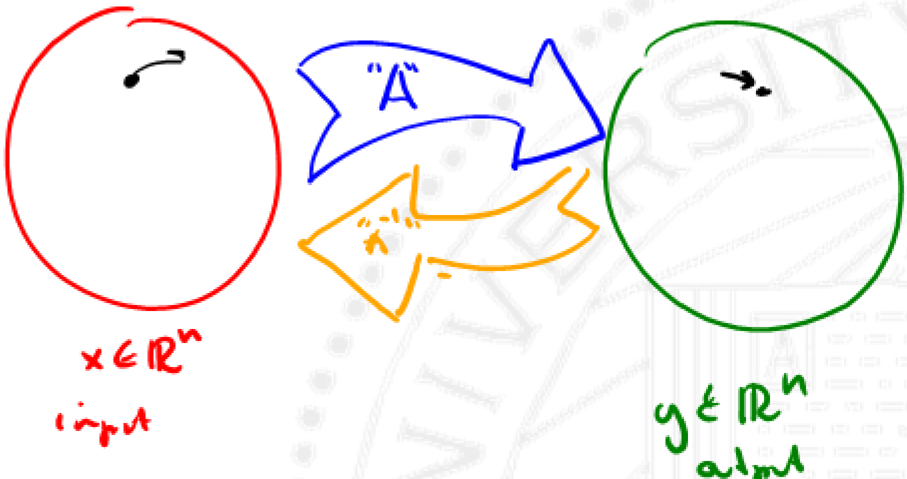
Square Matrices

Inverse

$\det(A) = \det(A^T)$

$\det(AB) = \det(A)\det(B)$

$A \in \mathbb{R}^{n \times n}$ invertible or non-singular if $\det(A) \neq 0$.



$N(A) = \{0\}$

"onto"

$R(A) = \mathbb{R}^n$

$\det(AA^T) = \det(A^TA) \neq 0$

A has a left & right inverse such that $AA^{-1} = A^{-1}A = I$ $\quad A^{-1} \in \mathbb{R}^{n \times n}$

- Columns of A are a basis for $\mathbb{R}^n$
  rows
- $y = Ax$ has a unique solution for every $y \in \mathbb{R}^n$

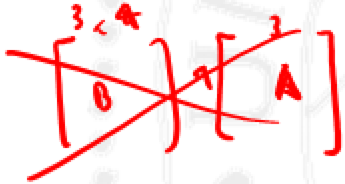
Interpretations of Inverse

$A \in \mathbb{R}^{n\times n}$ has an inverse $B = A^{-1} \in \mathbb{R}^{n\times n}$

- mapping associated w/ B undoes the mapping associated w/ A (pre- or post)

$x = By$ as a perfect (pre- or post) equalizer for the channel $y = Ax$.

$x = By$ is the unique solution to $y = Ax$.



Dual Basis Interpretation

Let a_i be a column of A $\quad \tilde{b}_i$ row of $B = A^{-1}$

$y = Ax$

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

$$x_i = \tilde{b}_i^T y$$

$$y = \sum_{i=1}^{n} (\tilde{b}_i^T y) a_i \quad \leftarrow \text{extracting } x_i\text{'s}$$

$\{\tilde{b}_1, \dots, \tilde{b}_n\}$ and $\{a_1, \dots a_n\}$ dual bases A

rows of B $\qquad$ columns of A

$$y = \sum_{i=1}^{n} x_i \underbrace{\tilde{b}_j^T a_i}_{\delta_{ij}} \qquad \delta_{ij} \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}$$

Rank of a Matrix

$A \in \mathbb{R}^{m \times n}$ $\operatorname{rank}(A) \hat{=} \dim(\underline{R(A)})$

all possible combinations of the columns of A

(non-trivial) facts:

$\operatorname{rank}(A) = \operatorname{rank}(A^T)$ ← not easy to prove.

$\operatorname{rank}(A)$ is the maximum number of independent columns or rows of A.

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

$\operatorname{rank}(A) + \dim(N(A)) = n$ ← Rank Nullity Theorem

Conservation of Dimension

$\boxed{\mathrm{rank}(A) + \dim(N(A)) = n}$:

→ $\mathrm{rank}(A)$ dimension of set "hit" by $y = Ax$.

– $\dim N(A)$ dimension of set of x crushed to $\{0\}$ by $y = Ax$

conservation of dimension: each dimension of input x is either crushed to $\{0\}$ or shows up in output.

n is # of degrees of freedom in x

$\dim N(A)$ is # of degrees of freedom lost in the mapping from x to $y = Ax$

$\mathrm{rank}(A)$ is # degrees of freedom in y.

$A = [\quad]^{10^6 \times 10^6} \quad [\quad]^{10^6 \times 1000} \, [\quad 10^6 \quad]^{1000}$

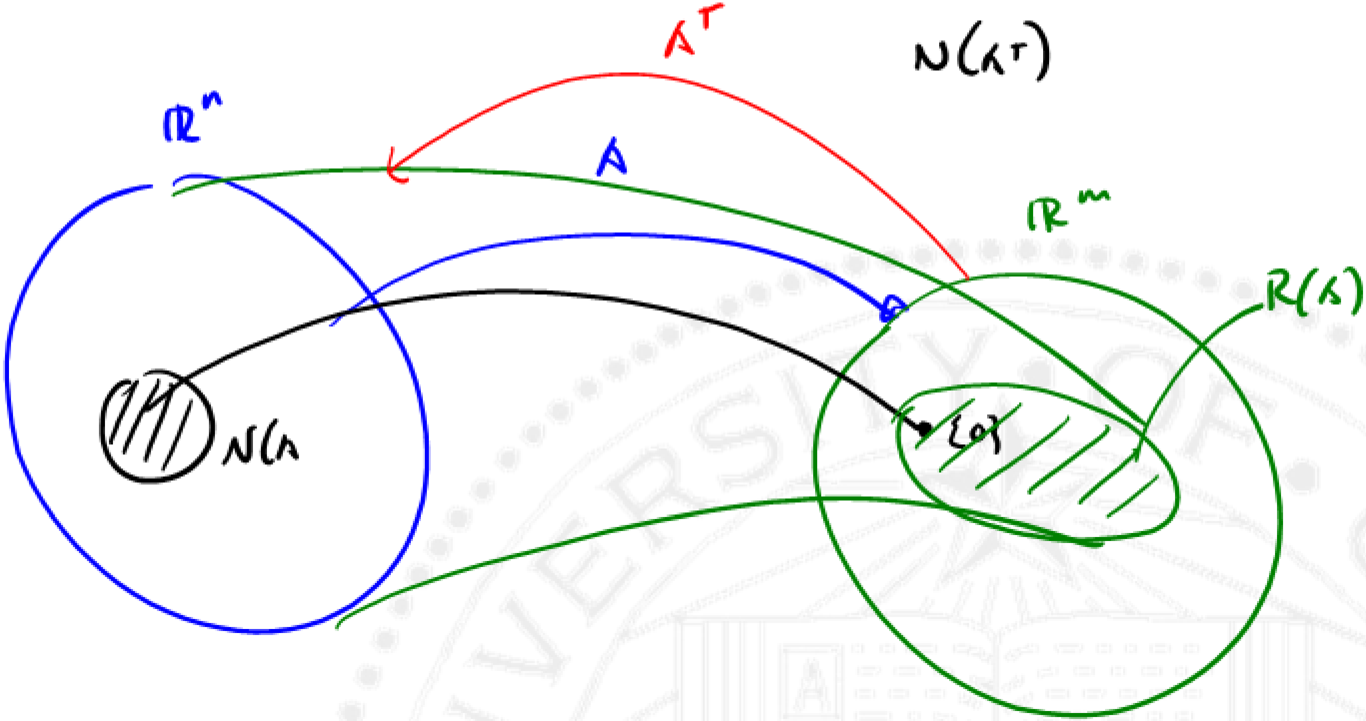
'Coding' Interpretation of Rank

$\text{rank}(BC) \leq \min(\text{rank}(B), \text{rank}(C))$

if $A = BC$ $\quad B \in \mathbb{R}^{m \times r}$ and $C \in \mathbb{R}^{r \times n}$ then

$\text{rank}(A) \leq r.$

$$x \xrightarrow{n} \boxed{A} \xrightarrow{m} y \quad \equiv \quad x \xrightarrow[n]{} \boxed{C} \xrightarrow{r} \boxed{B} \xrightarrow[m]{} y$$



$$N(A) \perp R(A^T)$$

$$\dim(R(A)) = \dim(R(A^T)) = \mathrm{rank}(A)$$
$$= \mathrm{rank}(A^T)$$

Application: fast matrix-vector multiplication

$y = Ax \qquad A \in \mathbb{R}^{m \times n} \qquad A = BC \quad B \in \mathbb{R}^{m \times r}$

$y = Ax$ — mn operations

$y = Ax = y = B(Cx) \rightarrow z = Cx \quad y = Bz.$

$rn + mr = (m+n)r$

$r \leq \min(m,n)$

Full Rank Matrices

$A \in \mathbb{R}^{m \times n}$

$\text{rank}(A) \leq \min(m, n)$

$= \underline{\text{FULL RANK}}$

For square matrices $m = n$, full rank means non-singular ($\det(A) \neq 0$, A^{-1} exists)

Both rows & columns are independent

For skinny matrices $\begin{bmatrix} | & | & | & | \end{bmatrix}$ columns are independent ($m \geq n$)

For fat matrices $\begin{bmatrix} — \\ — \\ — \\ — \end{bmatrix}$ rows are independent ($m \leq n$)

Change of Coordinates (1.3)

"Standard basis set" in $\mathbb{R}^n$: $(e_1, e_2 \dots e_n)$

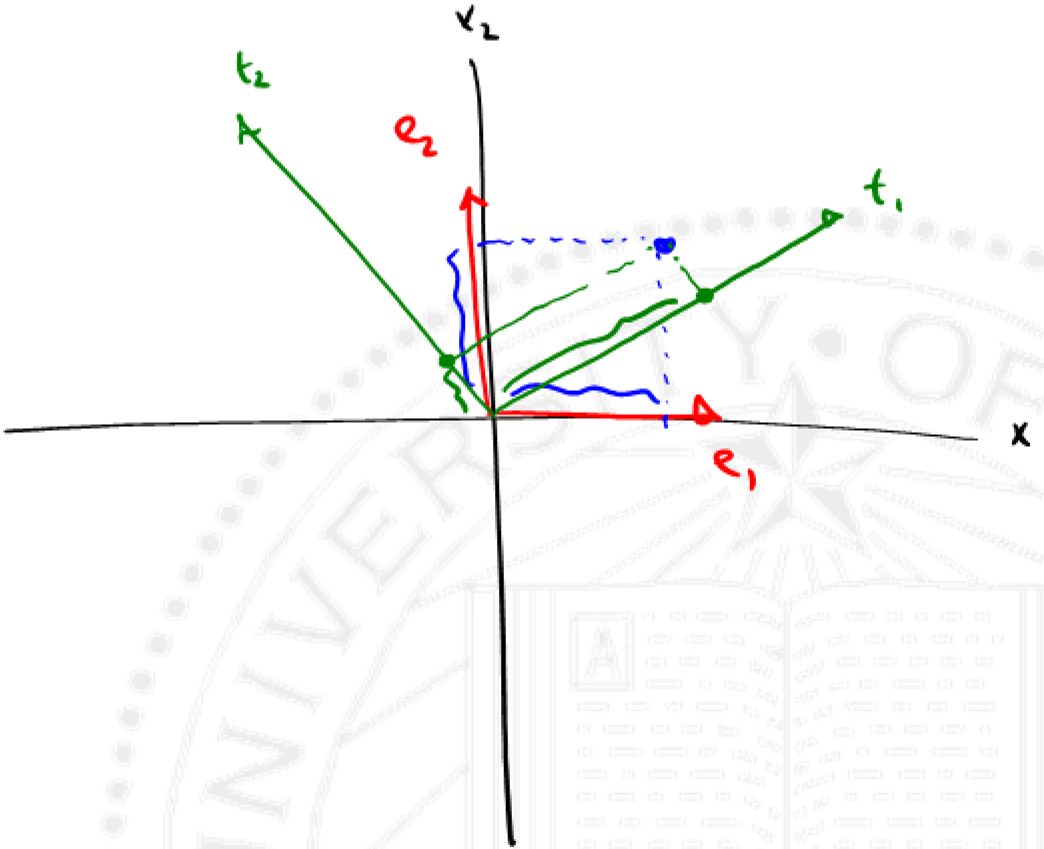
e_i: $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ ← ith posit e.g. $\mathbb{R}^3$: $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$$\underline{x} = \underbrace{\begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix}}_{I} \begin{bmatrix} x_1 \\ \vdots \\ x_2 \end{bmatrix} = I\underline{x}$$

$$\underline{x} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$$

basis $(t_1 \dots t_n)$ $\underline{x} = \tilde{x}_1 t_1 + \tilde{x}_2 t_2 + \dots \tilde{x}_n t_n = T\tilde{x}$

$$\boxed{\underline{\tilde{x}} = T^{-1}\underline{x}} \quad \text{or} \quad \boxed{\underline{x} = T\tilde{x}}$$



x_2
t_2
e_2
t_1
x_1
e_1

Change of Coordinates (2.3)

$x = T^{-1}\tilde{x}$ when $T = [t_1 \ldots t_n]$

$\tilde{x} = Tx$

T^{-1} transforms the standard basis coordinates into the t-coordinates

inner product with i^{th} row of T^{-1} extracts the t_i^{th} coordinate from $\underline{x}$.

Change of Coordinates (3.3)

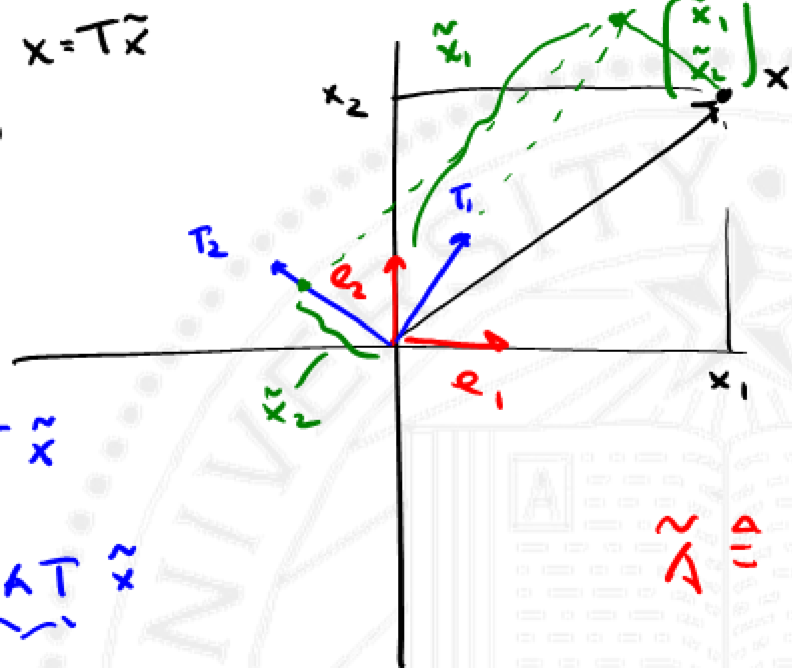
$y = \overset{n\times n}{A}x \qquad x = T\tilde{x}$

$y = AT\tilde{x}$

$y = T\tilde{y}$

$T^{-1}T\tilde{y} = T^{-1}AT\tilde{x}$

$\tilde{y} = \underbrace{T^{-1}AT}_{\text{similarity transform}}\tilde{x}$



$\begin{bmatrix}\tilde{x}_1 \\ \tilde{x}_2\end{bmatrix}_x = [t_1 \; t_2]^{-1}\begin{bmatrix}x_1 \\ x_2\end{bmatrix}$

$A, \tilde{A}$ | same rank, eigenvalues, eigenvectors, same det

$\tilde{A} \triangleq T^{-1}AT$

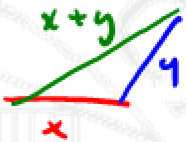
(Euclidean) norm

$\underline{x} \in \mathbb{R}^n$ $\qquad \|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{x^T x}$

$\|x\|_2$ 2-norm length of vector from origin

$\|\alpha x\| = |\alpha| \, \|x\|$ (homogeneity)

$\|x + y\| \le \|x\| + \|y\|$ (triangle inequality)



$\|x\| \geq 0$ (non-negativity)

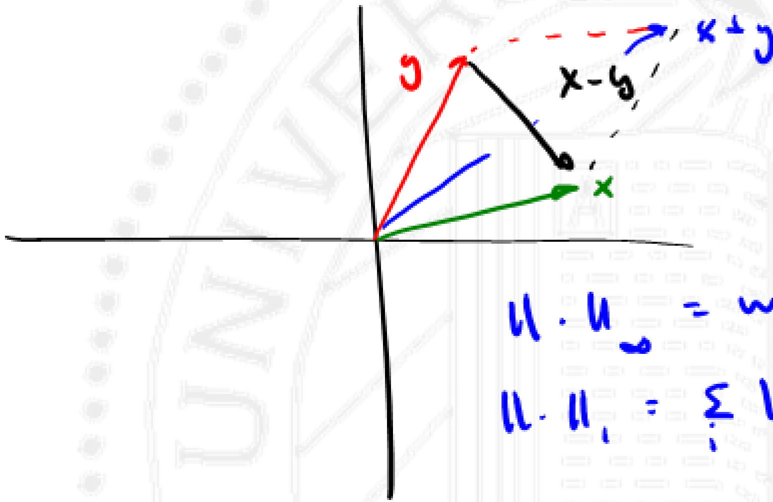
$\|x\| = 0 \rightarrow x = 0$ (definition)

RMS value and (Euclidean) distance

$$\mathrm{rms}(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{1/2} = \frac{\|x\|}{\sqrt{n}} \leftarrow \text{power}$$

$$\|x\|_2 \leftarrow \text{"energy"}$$

$$\mathrm{dist}(x,y) = \|x-y\|$$



$$\|\cdot\|_\infty = \max_i |x_i| \leftarrow \text{peak}$$

$$\|\cdot\|_1 = \sum_i |x_i|$$