Linear Algebra Review

Gabriel Hugh Elkaim

Linear Algebra Review

- Vector Spaces, subspaces
- Independence, basis, dimension
- Range, nullspace, rank
- Change of coordinates
  - Norm, angle, inner product.
- Norm, angle, inner product

Vector Spaces

- A *vector space* or *linear space* consists of:

a set $\mathcal{V}$ $\quad x+y$

a vector sum $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ $\quad +(x,y)$

a scalar multiplication: $\times : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$ $\quad \times(\alpha, v)$ $\quad \alpha v$

a distinguished element $0 \in \mathcal{V}$

Vector Space Properties

$x + y = y + x \quad \forall\, x, y \in \mathcal{V}$ (+ commutative)

$(x+y)+z = x+(y+z) \quad \forall\, x, y, z \in \mathcal{V}$ (+ associative)

$0 + x = x \quad \forall\, x \in \mathcal{V}$ (additive identity)

$\forall\, x \in \mathcal{V} \quad \exists (-x) \in \mathcal{V}$ such that $x + (-x) = 0$
(existence of an additive inverse)

$(\alpha\beta)x = \alpha(\beta x) \quad \forall\, \alpha, \beta \in \mathbb{R},\ x \in \mathcal{V}$

(scalar multiplication association)

$\alpha(x+y) = \alpha x + \alpha y$ (right distribution rule)

$(\alpha+\beta)x = \alpha x + \beta x$ (left distribution rule)

$1x = x, \quad \forall\, x \in \mathcal{V}$

Vector Space Examples

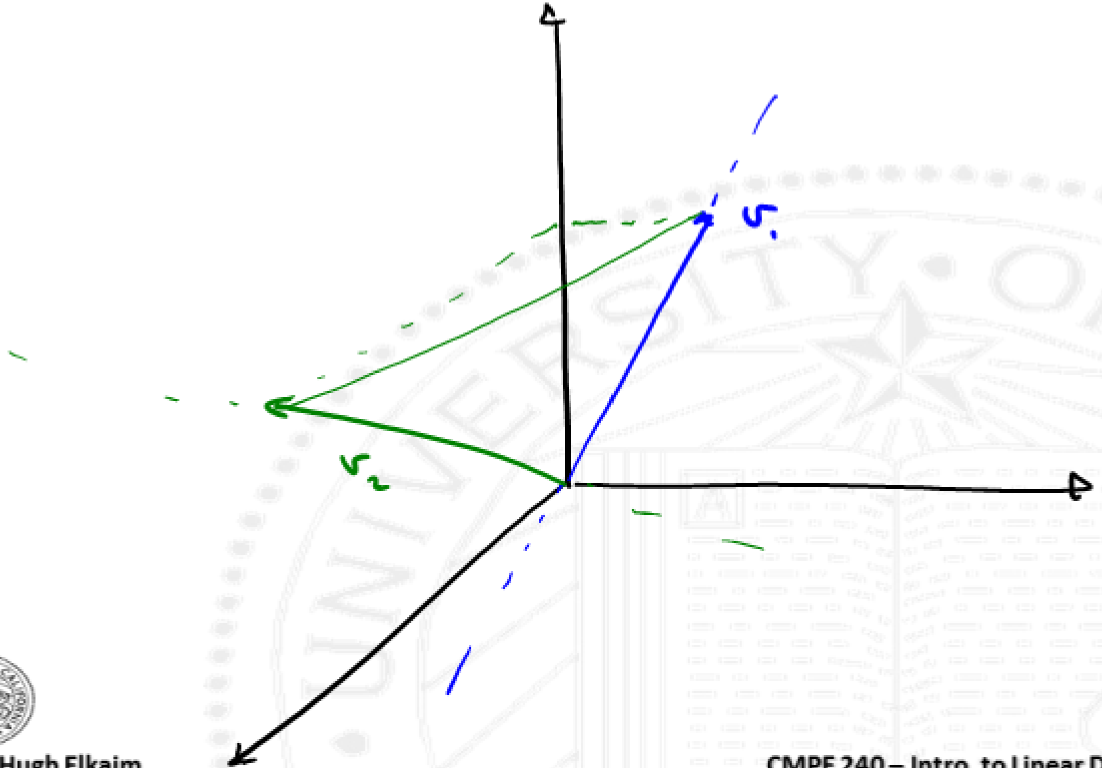
$V_1 = \mathbb{R}^n$ w/ standard elementwise vector addition and scalar multiplication

$V_2 = \{0\}$ where $\{0\} \in \mathbb{R}^n$ ← trivial

$V_3 = \text{span}\{v_1, v_2, \ldots v_k\}$ where

$\text{span}\{v_1, v_2 \ldots v_k\} \triangleq \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k \mid \alpha_i \in \mathbb{R}\}$

$v_1 \ldots v_k \in \mathbb{R}^n$



v_1
v_2

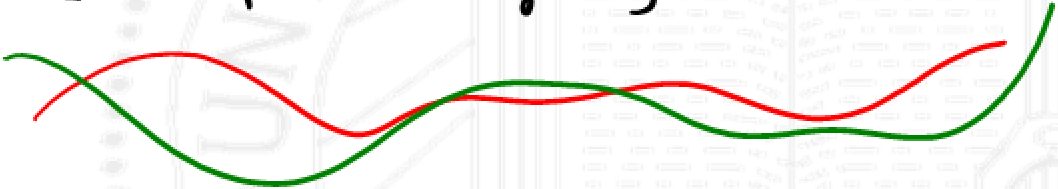
Vector Spaces of Functions

$V_d = \{ x : \mathbb{R}_+ \longrightarrow \mathbb{R}^n \mid x \text{ is differentiable} \}$ where vector sum is a sum of functions

$$(x + z)(t) = x(t) + z(t)$$

scalar multiplication $(\alpha x)(t) = \alpha x(t)$

a <u>point</u> in V_d is a trajectory in $\mathbb{R}^n$



Subspaces

A subspace of a vector space is a subset of a vector space which itself is also a vector space.

Subspace is closed under vector addition and scalar multiplication

V_1, V_2, V_3 are subspaces of $\mathbb{R}^n$

$V_5 = \{ x \in V_4 \mid \dot{x} = Ax \}$ points in V_5 are trajectories of $\dot{x} = Ax$.

Independent Set of Vectors

Property of a set of vectors $\{v_1 \dots v_k\}$ is independent

if: $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \phi \rightarrow \alpha_1 \dots \alpha_k = 0.$

No vector v_i can be expressed as a linear combination of the other vectors.

ex: $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \longrightarrow \cancel{\begin{bmatrix} 2 & 4 \\ & 4 \end{bmatrix}}$

$\alpha_1 v_1 + \dots + \alpha_k v_k = \beta_1 v_1 + \dots + \beta_k v_k \longrightarrow \alpha_1 = \beta_1 \dots \alpha_k = \beta_k$

Basis and Dimension

$\{v_1 \cdots v_k\}$ is a <u>basis</u> for a vector space V iff:

$\{v_1 \cdots v_k\}$ span V $\quad V = \text{span}\{v_1 \ldots v_k\}$

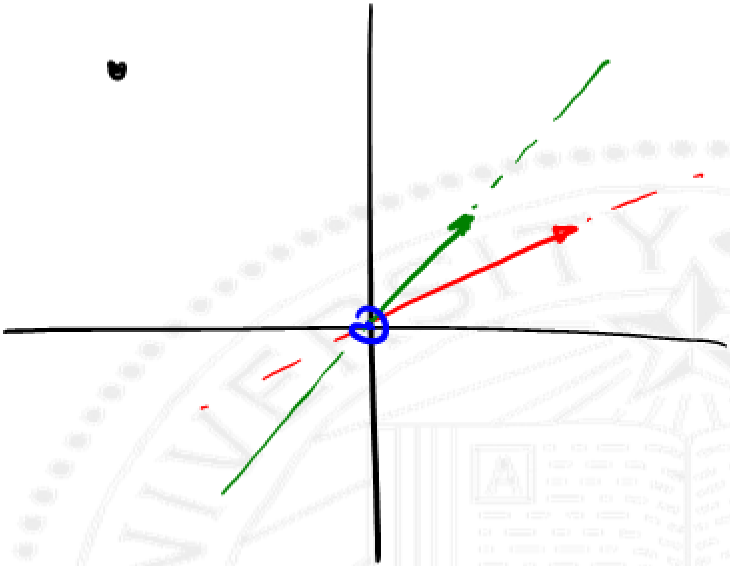
<u>and</u> $\{v_1 \ldots v_k\}$ is independent

every $v \in V$ can be uniquely expressed

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

$\text{Dim}(V) \triangleq$ # of v_k's that form a basis

"Cardinality"

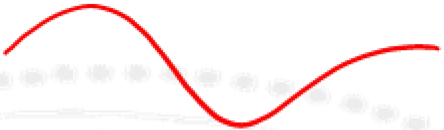


Jacobian

GRADIENT

$\underline{x} \longrightarrow \frac{\partial x}{\partial t}$

$\underline{x}(1)$



$\frac{\partial}{\partial t}\underline{x} \longrightarrow D\underline{x}$

"Jacobian"

Nullspace of a Matrix

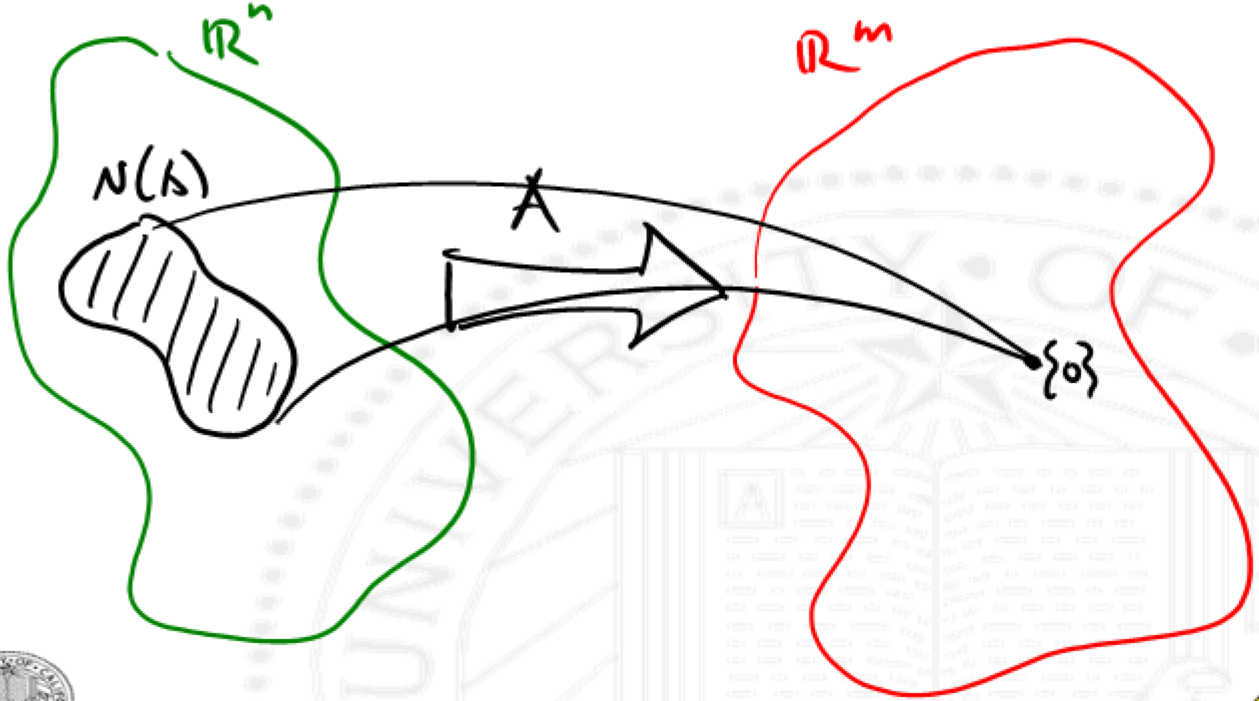
$A \in \mathbb{R}^{m \times n}$

null space
"kernel"

$$\mathcal{N}(A) \triangleq \{x \in \mathbb{R}^n \mid Ax = \overset{\mathbb{R}^m}{\overset{\downarrow}{0}}\}$$

$\mathcal{N}(A)$ is a set of vectors mapped to zero by $y = Ax$

$\mathcal{N}(A)$ is a set of vectors that are orthogonal to all rows of A.



$\mathbb{R}^n$
$\mathbb{R}^m$
N(A)
A
{0}

Zero Nullspace

$\mathbb{R}^n$

$N(A) = \{0\}$ — "one to one"

$y = Ax$

$x = By = BAx = Ix = x$

$\underline{x}$ can always be determined from $y = Ax$

$y = Ax$ transformation loses no information

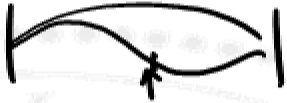
map from $x \to y$ different x's map to different y's.

$\det(A^T A) \neq 0$. $\{$ perfect decoder

A has a LEFT inverse $B \in \mathbb{R}^{n \times m}$ such that $BA = I$.

Interpretations of Nullspace

$y = Ax$ $\qquad z \in N(A)$



$y = Ax$ is a measurement $\longrightarrow$ z is undetectable by any sensors

x and $x+z$ are indistinguishable

$N(A)$ is the ambiguity in x from measurement $y = Ax$.

$y = Ax$ is the output from input x.

z is input w/ no result

$x+z$ and x have the same result.

$N(A)$ freedom of choice

Range of a Matrix

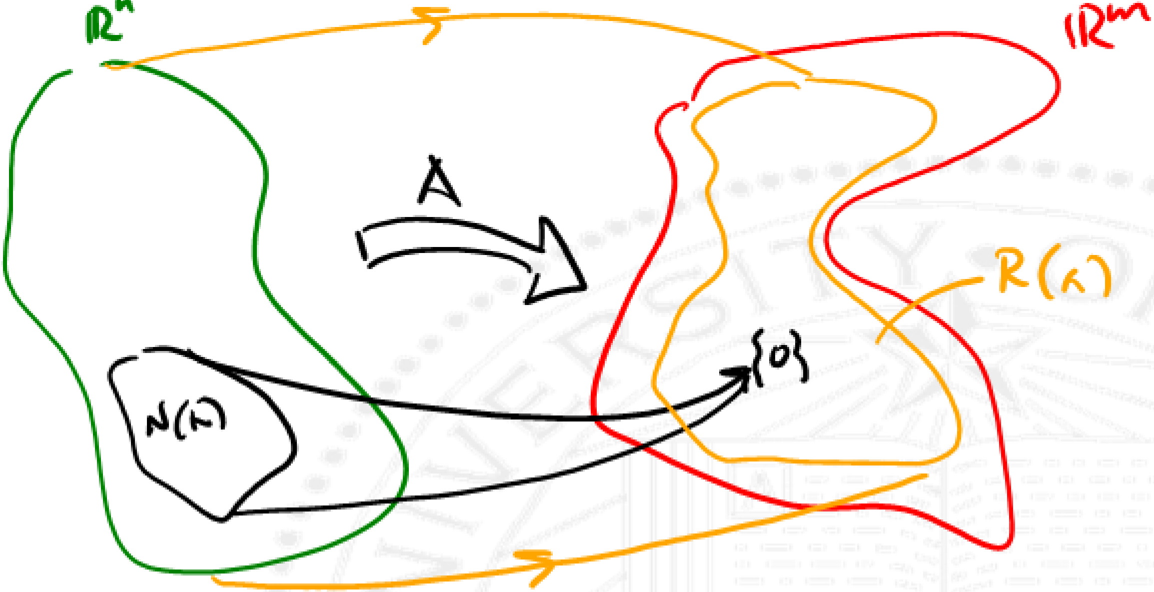
$R(A)$: range of A

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

Set of all values that can be "hit" by $y = Ax$

Span $\{a_1, a_2, \ldots, a_n\}$

Set of vectors $\{y\}$ for which $y = Ax$ can be solved.



$\mathbb{R}^n$
$\mathbb{R}^m$
A
N(A)
R(A)
{0}

"*Onto*" Matrices

$y = Ax \quad x = B y_m$

$y = AB y_m = I y_m$

$y = y_m$

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

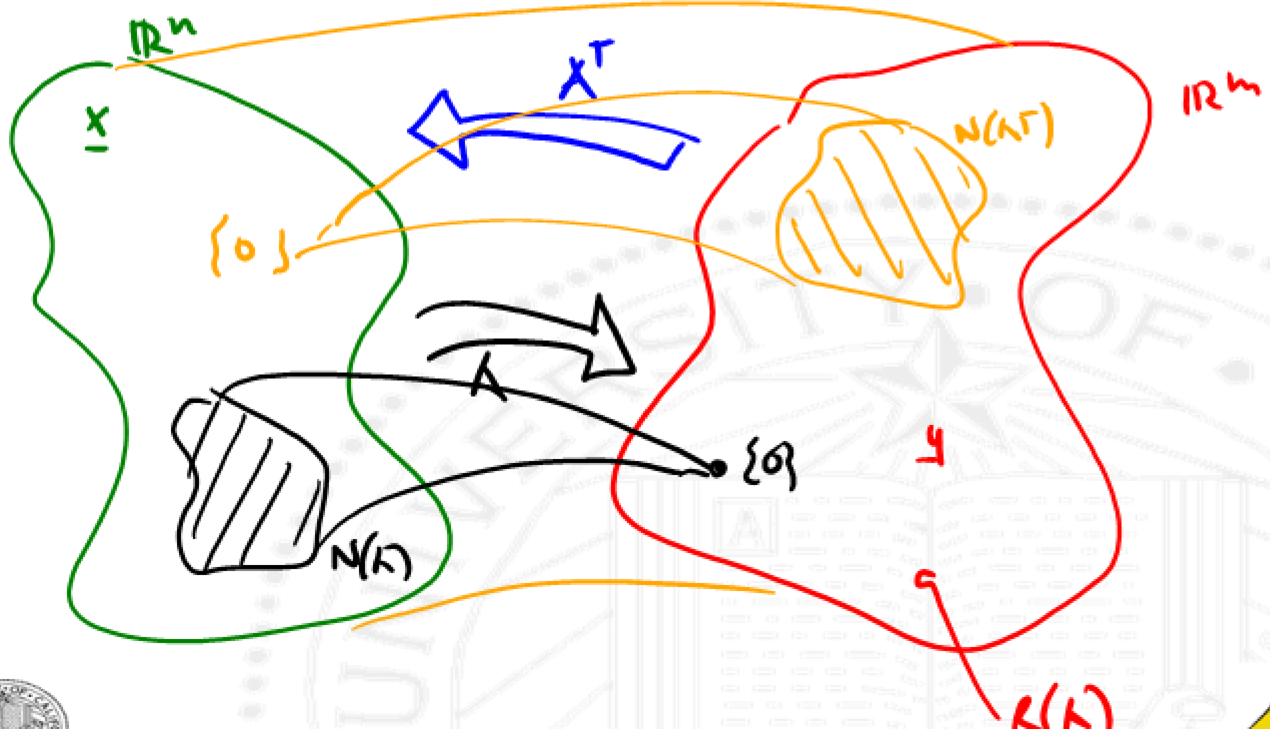
$y = Ax$ can be solved for any $\{y\}$.

Columns $\{a_1, a_2, \ldots, a_n\}$ span $\mathbb{R}^m$

A has a RIGHT inverse $B \in \mathbb{R}^{n \times m}$ $\quad AB = I$

rows of A are independent $\{\tilde{a}_i^T\}$

$N(A^T) = \{0\}$ and $(AA^T) \neq 0$.



$\mathbb{R}^n$
$\underline{x}$
A^T
$\mathbb{R}^m$
$N(A^T)$
$\{0\}$
A
$N(A)$
$\{0\}$
$\underline{y}$
$R(A)$

Interpretations of Range

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

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

$R(A) = \{\underbrace{Ax}_{\text{image of } A} \mid x \in \mathbb{R}^n\}$

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

$y = Ax$ or measurement

$y = v$ possible or consistent sensor measurement.

$y = w$ impossible sensor measurement / inconsistent signal

$y = Ax$ is an output/result

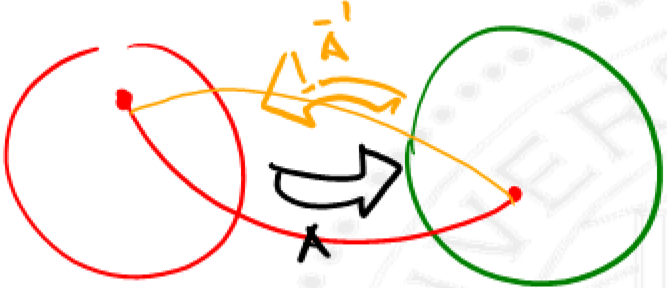
$y = v$ is a possible result or output

$y = w$ is an impossible result or output

Square matrices $(n=m)$ **Inverse**

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

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



$x \in \mathbb{R}^n$
input

$y \in \mathbb{R}^n$
output

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

"onto"

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

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

A has both a left and right inverse such that $\underline{A^{-1}A = AA^{-1} = I}$.

Columns of $A \rightarrow \{a_1, \ldots, a_n\}$ are a <u>basis</u> for $\mathbb{R}^n$

$y = Ax$ 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 with 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 <u>unique</u> solution to $y = Ax$

Dual Basis Interpretation

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

$$y = Ax = a_1x_1 + a_2x_2 + \dots + a_nx_n \qquad x_i = \tilde{b}_i^T y.$$

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

$\{\tilde{b}_1 \dots \tilde{b}_n\}$ and $\{a_1 \dots a_n\}$ dual basis for A.

rows of B columns of A

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

Rank of a Matrix

$A \in \mathbb{R}^{m \times n}$ $\quad \text{rank}(A) \triangleq \dim(\underline{R(A)})$

all possible combination of columns of A

(non-trivial) facts

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

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

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

RANK NULLITY THEOREM

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

Conservation of Dimension

$\underline{\text{rank}(A) + \dim(N(A)) = n}$:

→ rank(A) dimension of set "hit" by $y = Ax$

→ dim (N(A)) dimension of set x that are crushed to {0} by $y = Ax$

conservation of dimension: each dimension of input (x) is either crushed to {0} or shows up in the output (y).

n is the # of degrees of freedom in x

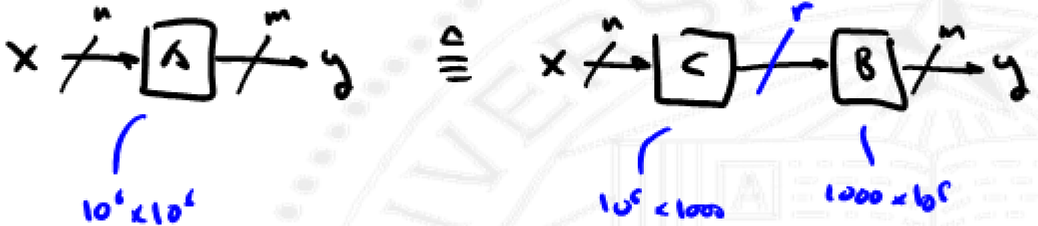
dim (N(A)) is # of degrees of freedom lost in $y = Ax$.

rank(A) # of degrees of freedom in y.

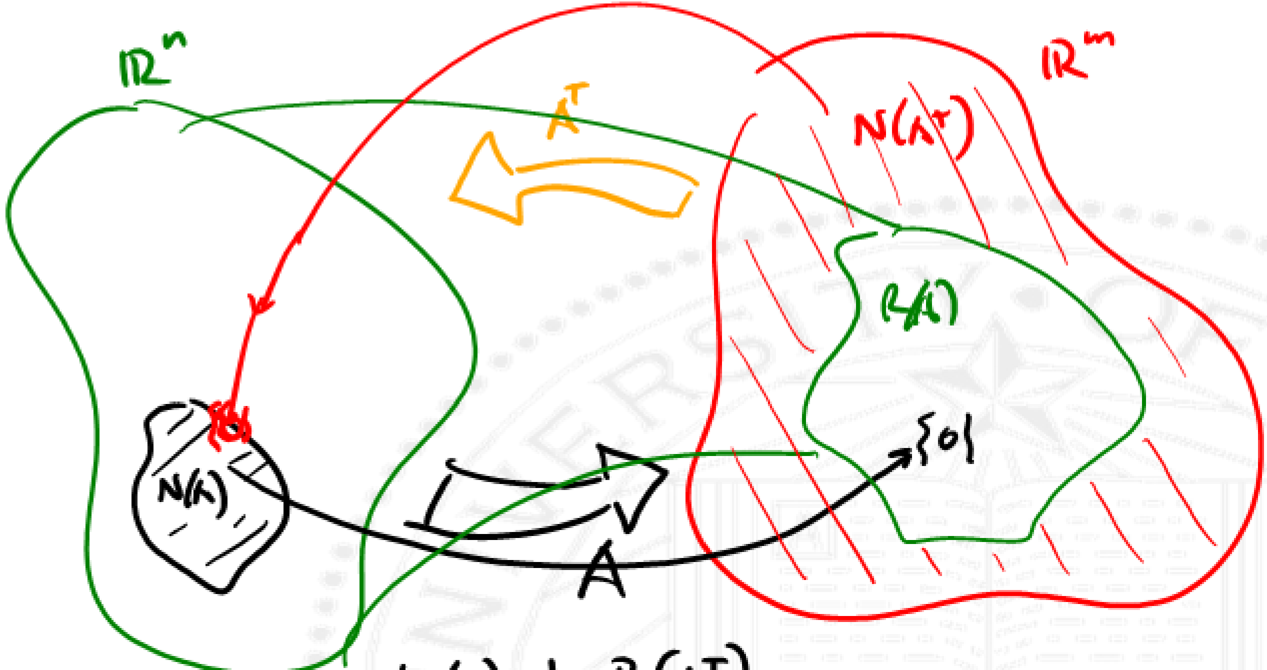
'Coding' Interpretation of Rank

$\mathrm{rank}(BC) \le \min(\mathrm{rank}(B), \mathrm{rank}(C))$

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



$\mathrm{rank}(A) \le r.$



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

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

Application: fast matrix-vector multiplication

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

$y = Ax \leftarrow m \cdot n$ operations

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

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

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

Full Rank Matrices

$A \in \mathbb{R}^{m \times n}$ $\quad \mathrm{rank}(A) \le \min(m,n)$

$= $ <u>FULL RANK</u>

for <u>square matrices</u> $(m=n)$, full rank means non-singular

$\det(A) \neq 0$, A^{-1} exists

Both rows & columns of A are independent

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

for fat matrices $\begin{bmatrix} \text{---} \\ \text{---} \\ \text{---} \end{bmatrix}$ rows are independent $(m \le n)$

Change of Coordinates (1.3)

"Standard Basis Set" in $\mathbb{R}^n \longrightarrow \{e_1, e_2, \dots, e_n\}$

$e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th}$ position eg: $\mathbb{R}^3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

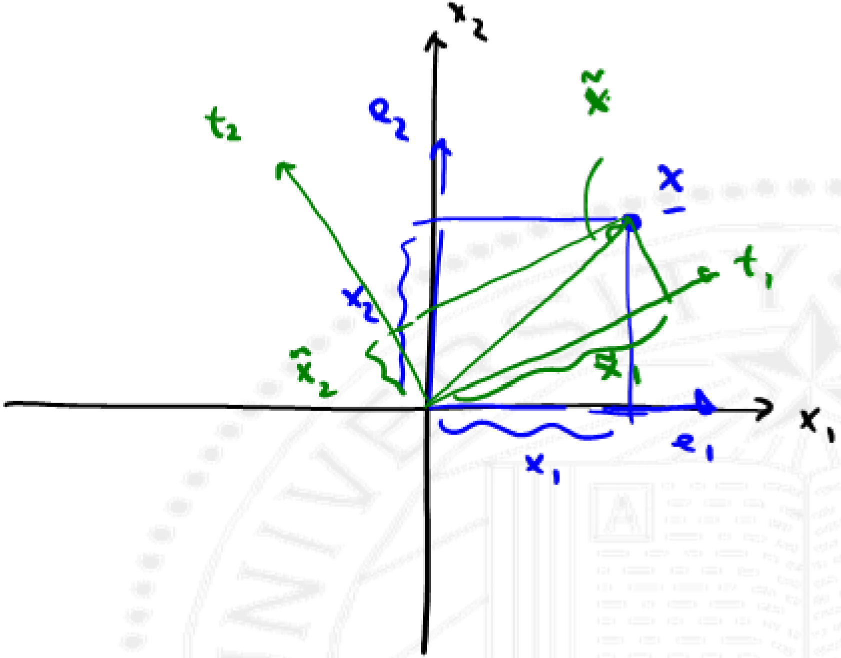
$$\underline{x} = \underbrace{\begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix}}_{I} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = I\underline{x}$$

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

basis $(t_1 \dots t_n)$

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

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



x_2
e_2
t_2
$\tilde{x}$
$\underline{x}$
t_1
x_2
$\tilde{x}_2$
$\tilde{x}_1$
x_1
e_1
x_1

Change of Coordinates (2.3)

$\underline{x} = T\,\underline{\tilde{x}}$ where $T = \{t_1, t_2, \ldots, t_n\}$

$\underline{\tilde{x}} = T^{-1}\underline{x}$

T^{-1} transforms the standard basis at 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)

$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

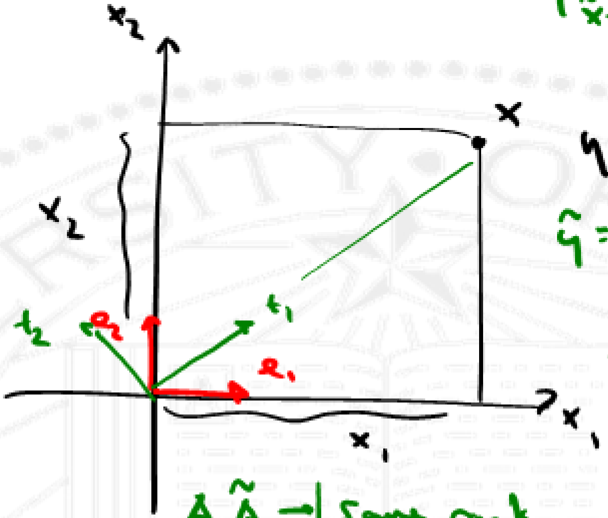
$y = \underbrace{A}_{\mathbb{R}^{m \times n}} x \qquad x = T\tilde{x}$

$y = AT\tilde{x}$

$y = T\tilde{y}$

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

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



$y = Ax$

$\tilde{y} = \tilde{A}\tilde{x}$

$\uparrow$

$T^{-1}AT$

$A, \tilde{A} \rightarrow$ same rank, eigenvalues, eigenvectors, same determinant

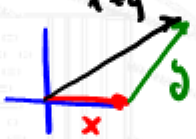
(Euclidean) norm

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

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

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

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



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

$\|x\| = 0 \longrightarrow x = 0$ (definiteness)

RMS value and (Euclidean) distance

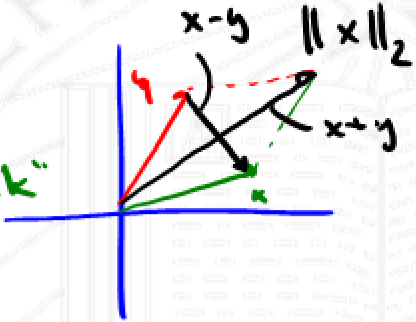
$$\mathrm{rms}(x) \triangleq \left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right]^{1/2} = \frac{\|x\|}{\sqrt{n}} \leftarrow \text{"power"}$$

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

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

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

$$\|\cdot\|_1 \triangleq \sum_i |x_i|$$

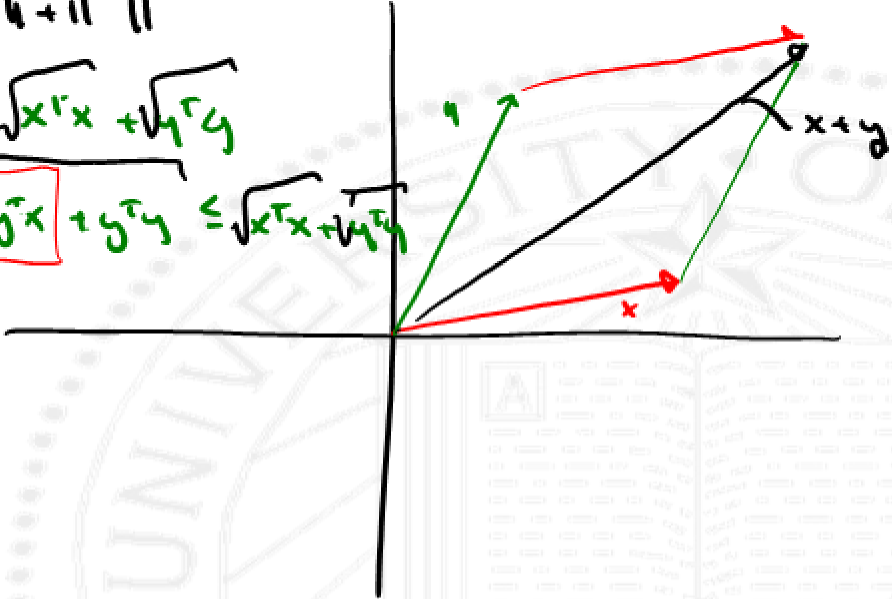


TRIANGLE INEQUALITY

$$\|x+y\| \le \|x\| + \|\ \|$$

$$\sqrt{(x+y)^T(x+y)} \le \sqrt{x^Tx} + \sqrt{y^Ty}$$

$$\sqrt{x^Tx + \boxed{x^Ty + y^Tx} + y^Ty} \le \sqrt{x^Tx} + \sqrt{y^Ty}$$



$\langle x|y\rangle$ → "bracket" $|x\rangle\langle y|$ — "ketbra"

Inner Product ← dot product

$\langle x,y\rangle := x_1y_1 + x_2y_2 + \dots + x_ny_n = x^Ty$

$\left.\begin{array}{l}\langle \alpha x,y\rangle = \alpha\langle x,y\rangle \\ \langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle\end{array}\right\}$ LINEAR

$\langle x,y\rangle = \langle y,x\rangle$ reflexivity

$\langle x,x\rangle \geq 0$ positivity

$\langle x,x\rangle = 0 \leftrightarrow x=0$ "definite"

$f(y) = \langle x,y\rangle$ lin func $\mathbb{R}^n \to \mathbb{R}$

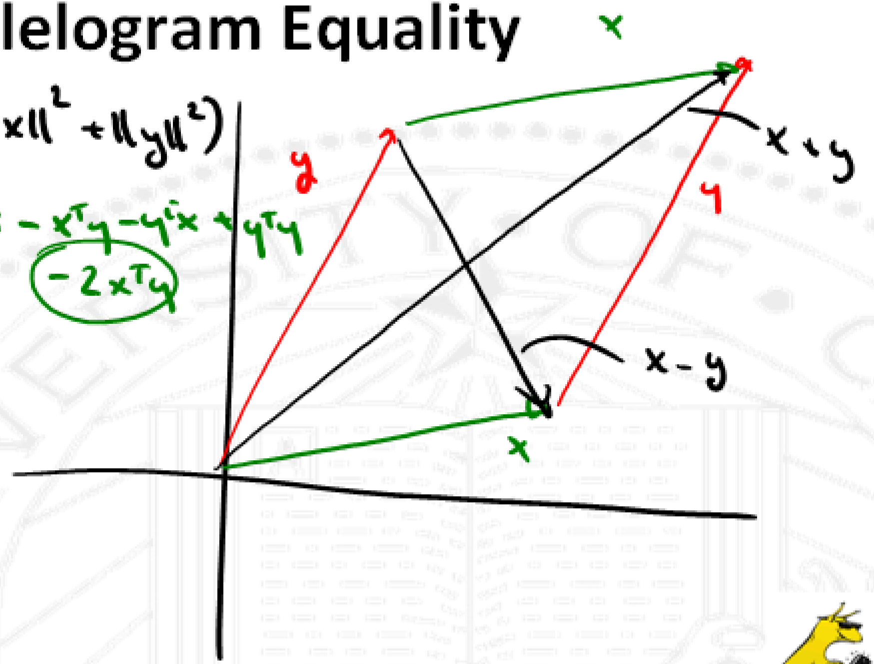
$f(y) = \underbrace{x^T}_{A \in \mathbb{R}^{1\times n}} y$

Parallelogram Equality

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

$x^Tx + x^Ty + y^Tx + y^Ty$ $\quad$ $x^Tx - x^Ty - y^Tx + y^Ty$

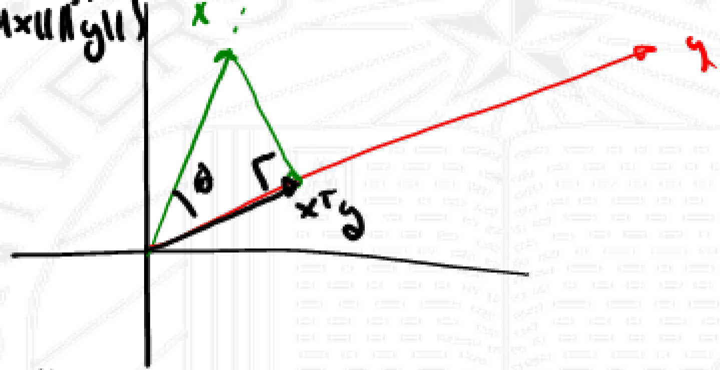
$2x^Ty$ $\quad$ $-2x^Ty$



Cauchy-Schwartz Inequality

$x, y \in \mathbb{R}^n$ $\qquad |x^T y| \leq \|x\| \|y\|$

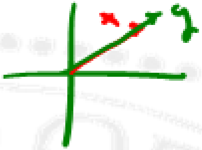
$\theta \triangleq \angle(x, y) = \cos^{-1}\left\{\frac{x^T y}{\|x\|\|y\|}\right\}$



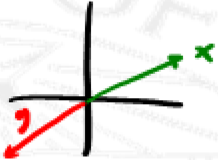
$x^T y = \|x\| \cdot \|y\| \cos\theta$

Special cases of Cauchy-Schwartz

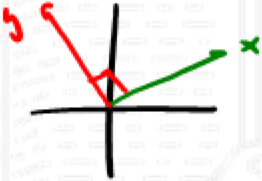
x, y aligned $\quad \theta = 0 \quad x^T y = \|x\|\|y\|$



x, y opposed $\quad \theta = \pi \quad x^T y = -\|x\|\|y\|$

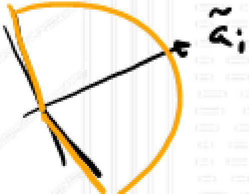


$x \perp y$ perpendicular $\quad \theta = \frac{\pi}{2} \quad x^T y = 0$



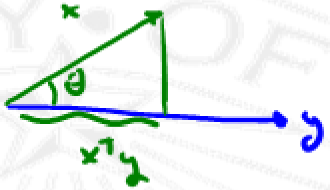
$y = Ax$ — all y_i's are ≥ 0.

$\langle \tilde{a}_i, x \rangle$

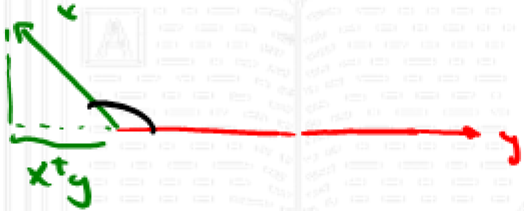


Half-space Interpretation

$x^T y > 0$ means $\angle(x,y)$ acute $\quad \angle\theta < 90^\circ$

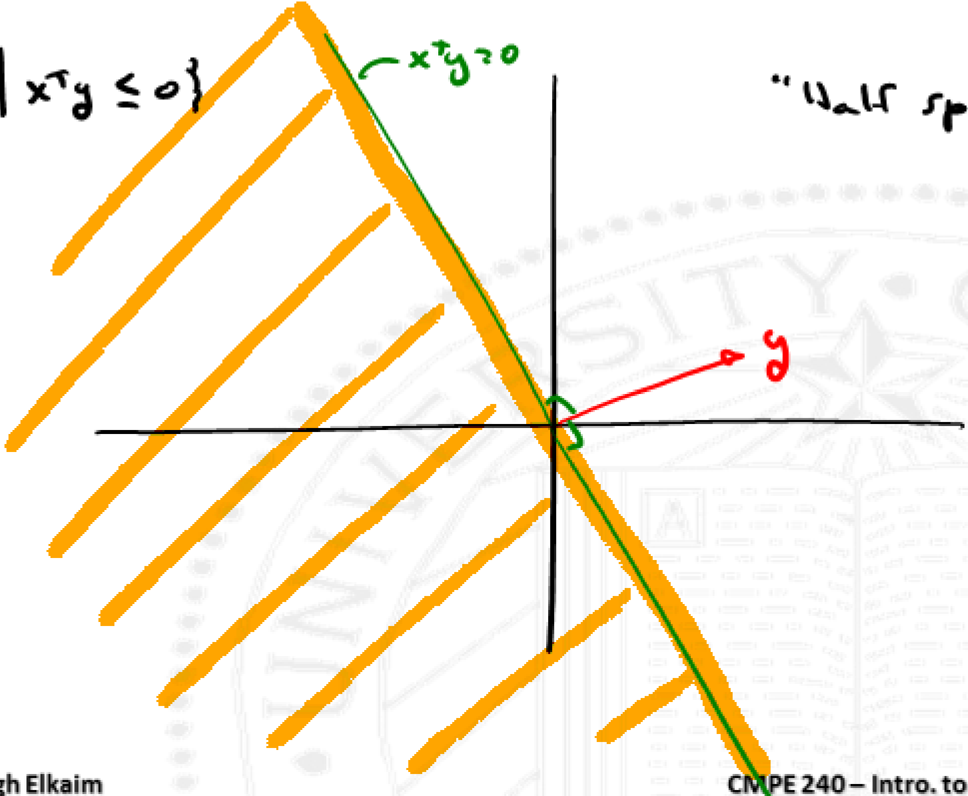


$x^T y < 0$ means $\angle(x,y)$ obtuse

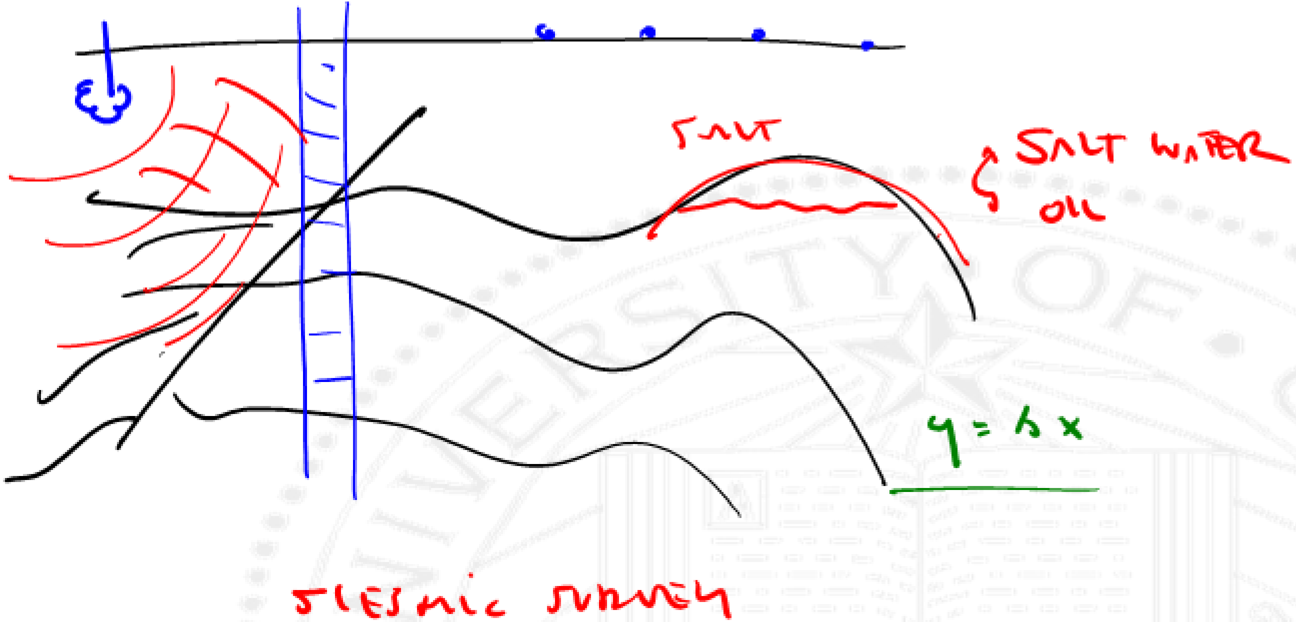


$\{ x \mid x^T y \le 0 \}$

"Half space"



Questions?


SALT
SALT WATER
OIL
y = A x
SEISMIC SURVEY