SDS 387 Linear Models
Fall 2024
Lecture 11 - Tue, Oct 1, 2024
Instructor: Prof. Ale Rinaldo
Some useful references:
<ul> <li>Mostrix Analysis by Horn &amp; Johson } Very regorous &amp; comprehensive</li> <li>Mostrix Analysis by Bhatia</li> <li>Mostrix Computations by Golub } Algorithmic focus</li> <li>Mostrix Perturbation Theory by Sun &amp; stewart</li> </ul>
<ul> <li>Linear Algebre Done Right by Axler</li> <li>Introduction to Applied Linear Algebre by Boyd available online</li> </ul>
- Appendix to Plane Answers to Simple Questions (available online by Christenses from Springerline)
. For the statistics /ML results about linear models, we will use
next the book. Learning Theory from First Principles
by Brancis Bach (available online)
$\mathcal{O}$

E LINEAR ALGEBRA RECAP
. We will be working in IR? but much of what we say
holds in more general spaces
· A vector space (over R): a set closed wit scalar
nuttiplication and addition. M vector space
It has a zero element. ZEM -> aze M Sack
at Q = 2 V zeus element ary EM -> zeg EM
· A linear subspace N of M is a subset that is
oulso a vector space, <u>Example</u> : In Rª,
$\begin{cases} \mathcal{X} = \begin{bmatrix} \mathcal{Z}_{i} \\ \vdots \\ \vdots \\ z_{d} \end{bmatrix};  \mathcal{X}_{k+i} = \cdots = \mathcal{X}_{d} = 0 \end{cases}$
In IR? a linear subspaces are lines through the origins
· A (funite) subset of M [v.,, vel] is
a set of linearly interpendent vectors or points
$if \qquad \qquad$
lunear combination
· A set of linearly independent vectors
{v1,, vk} spows a subspace N of M
when every see N can be written as a
linear combination of the his.

	In this case, §0,, vu? is called a basis of N. Boses are not unique, but the number of elements in each basis is the same
· · · · · · · · · · · ·	and it is called the dimension or rank of the subspace.
	Forts: if [Vi,, Vh] is a bosis for N then the N Z   a,, are ell there unique
	s.t. $x = \sum_{i=1}^{r} \alpha_i v_i$ If $N_1$ and $N_2$ are subspaces, so is
	$\mathcal{N}_{i} \neq \mathcal{N}_{2} = \{ \boldsymbol{x} : \boldsymbol{x} = \boldsymbol{x}_{i} \neq \boldsymbol{z}_{2}  \boldsymbol{x}_{i} \in \mathcal{N}_{i} \; \boldsymbol{x}_{2} \in \mathcal{N}_{2} \}$ $\mathcal{N}_{i}  \mathcal{N}_{2}$
· · · · · · · ·	What about N. U N2 ? No
· · · · · · · · ·	$H = N_1 \cap N_2 = O  \text{then} \\ rounk (N_1 + N_2) = rounk(N_0) + rounk (N_2)$
	In $\mathbb{R}^{d}$ (and in many other space) we have on unner product, a function on $\mathbb{R}^{d} \times \mathbb{R}^{d}$ that is symmetric ( <xcy>=<y,2>), linear (3)</y,2></xcy>

$(\langle \alpha \alpha, \beta \gamma \rangle = \alpha \beta \langle \alpha, \gamma \rangle  \alpha, \beta \in \mathbb{R})$
and positive definite <2,2>>0
$\ln R^{d} = \begin{bmatrix} x_{1} \\ \vdots \\ z_{d} \end{bmatrix} = \begin{bmatrix} y_{r} \\ \vdots \\ y_{d} \end{bmatrix}$
$\langle n, y \rangle = \chi^{T} y = y^{T} x = \underset{\chi = 1}{\swarrow} \pi_{\chi} y_{\chi}$
This gives the Euclidean norm 11211=/{a.x>
. On the some space you can define more than one
inner product. (Example: 17 2120
$= \underbrace{\overrightarrow{x}}_{i} \underbrace{x}_{i} \underbrace{y}_{j} \underbrace{\overrightarrow{x}}_{i,j}$
$\tilde{\sim}, \delta$
Inner products ablow to define orthogonality;
. Inner products able to define orthogonality: $x$ and $y$ are orthogonal when $\langle x_{cy} \rangle = 0$
Inner products allow to define orthogonality: $x$ and $y$ are orthogonal when $\langle x_{cy} \rangle = 0$
An orthogonal basis is a basis consisting of
Inner products ablow to define orthogonality: x and y are orthogonal when <2cy>=0 An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal
Inner products allow to define orthogonality: x and y are orthogonal when <2cy>=0 An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have
Inner products allow to define orthogonality: x and y are orthogonal when <2cy>=0 An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product)
Inner products able to define <u>orthogonality</u> ; 2 and y are <u>orthogonal</u> when $(2x,y) = 0$ An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product).
Inner products allow to define <u>orthogonality</u> ; x and y are <u>orthogonal</u> when $\langle x, y \rangle = 0$ An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product).
Inner products allow to define <u>orthogonality</u> ; x and y are <u>orthogonal</u> when $\langle x_{cy} \rangle = 0$ An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product). If vi,, vie is a basis for some subspace
<ul> <li>Inner products allow to define <u>orthogonality</u>:</li> <li>2 and y are <u>orthogonal</u> when &lt;2003&gt;=0</li> <li>An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product).</li> <li>If vi,, ve as a basis for some subspace there always exists on orthonormal basis that</li> </ul>
<ul> <li>Inner products allow to define <u>orthogonality</u>:</li> <li>x and y are <u>orthogonal</u> when <x cy=""> = 0</x></li> <li>An orthogonal basis is a basis consisting of orthogonal vector. An the basis is orthonormal if it is orthogonal and all its elements have unit norm (the norm induced by inner product).</li> <li>If vs,, ve is a basis for some subspace there always exists on orthonormal basis that an orthogonal basis or orthonormal basis that can be constructed using vs,, ve. This process</li> </ul>

is known as Gran-Schnidt arthogonalization)
$y_{i} = \frac{v_{i}}{h v_{i}} $
for $n=2, \dots, k$ let $\int W_n = v_n - \frac{\xi}{j=r} \langle v_n, w_j \rangle w_j$
$   \left( \begin{array}{c} y \lambda = \frac{W \lambda}{\ W \lambda\ } \\ \end{array} \right) $
Then y',, ye are an orthonormal 00-315 HW
or of S is a subspace of M, the orthogonal
$S' = \{x \in M : \langle x_i y \rangle = 0 \; \forall y \in S \}$
$F_{\text{ext}}$ $S \cap S = \{0\}$
La Any vector z e M can be written uniquely as
$x = x_s + x_{s^{\pm}}$ where $x_s \in S$ direct $x_{s^{\pm}} \in S^{\pm}$ sum of course $\langle x_s, x_{s^{\pm}} \rangle = 0$
As a result M = S + S to and
$Four (M) = Four (S) + Four (S^{\perp})$
$\frac{Foct}{S_{c}};  lf  S_{c}  ound  S_{c}  ove  subspaces$ $\left(S_{c} \cap S_{c}\right)^{+} = S_{c}^{+} + S_{c}^{+}$ (5)

	MARICES
· · · · · ·	In Rd e vector is a 1-don array. A matrix is a 2-olim array:
  	$A = (A_{i}) ) i = (, n) rows \in \mathbb{R}^{n} \times \mathbb{R}^{n}$ $j = (, n) columns$
· · · · · · ·	set of notrices is closed with scalar multiplications and addition (if the matrices have some size)
	Nation of product: $A = B = C \qquad C_{n, \bar{n}} = \sum_{e=i}^{n} A_{i, e} B_{e_{n, \bar{n}}}$ $M = M = M = M = M$
.     .     .     .       .     .     .     .       .     .     .     .       .     .     .     .       .     .     .     .       .     .     .     .       .     .     .     .       .     .     .     .	Big issue: non commutativity. In general ABZBA
· · · · · ·	A R(A): linear subspace of R spanned mxn by columns of A
.       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .         .       .       .       .	Kernel (AS) lineor subspace of 1R <sup>n</sup> of null space (A) the form french Az=0 }
· · · · · ·	Ani oth column of A

· · · · · · · ·	The trows pose of $A = (Anis)$ is the metric
	$A^{T} = (A_{j,i})^{n \times n} \qquad Na^{te} \cdot (AB)^{T} = B^{T}A^{T}$
· · · · · · · · ·	A is square when m=n. A square matrix is mxn
	dragonal when all elements Arijsos Fizi
· · · · · · ·	I I a diagonal notions with Idii = 1 thi
	$I_n A = A T_u = A$
· · · · · · · · ·	The inverse of A is the matrix A' sit.
· · · · · · · ·	$A^{-i}A = AA^{-i} = I_n$ unique
· · · · · · · ·	$\frac{N_0 t_{e}}{\Delta B} = B' A''$
• • • • • • • •	if A has an inverse it is said to be given in an - ringular
· · · · · · · · · ·	This happens $r(f)$ rouce $(A) = \eta$
· · · · · · · ·	round spece the of hineary independent country or rous
· · · · · · · · ·	$n \times n$ $vane (null (A)) = n - v$
• • • • • • • •	A metrix U is orthogonal when its columns

ove orthonormal vectors. Then	
$ \begin{array}{c} \cdot \cdot$	
$\cdots \cdots $	• •
	• •
	• •
· Trace of square matrix A is Tr(A) = 2 Ain	• •
Tr() a linear function : Tr (aA+B) =	
$f_{1}(\alpha)$	
~ TRA/ FIRES	
tr(.) has cyclic property	• •
(A + B + A + A + A + A + A + A + A + B + C) = (A + B + C) = (A + A + B + C) = (A + A + A + A + A + A + A + A + A + A	• •
$(\Delta C R)$	
$= \frac{1}{2} \left[ \frac{1}{2} \left[ \frac{1}{2} \right] \right]$	
In fact you can define an inner product	
	• •
over square matrices	• •
$\langle A B \rangle = th (AB)$	
SPECIFAL PROPER NET A CARACTER AND A	• •
Let A Number 2 is an eigenvalue of	A
A corresponding eigenvertor is a vector x=R	ŝ.
$A_{\mathcal{X}} = \lambda_{\mathcal{X}} \qquad \left( \left( A - \lambda \Gamma \right)_{\mathcal{X}} = 0 \right)$	
$\mathcal{L} \neq \bigcirc $	
	• •
8)	) [

<ul> <li>We can find all the eigenvalues of A by solving The polynomial equation det (A - II) = 0</li> <li>If A has represented to A,, he than a representative di,, he than a representative di,, he than a representative distribution of the subspace null (A - dIn) is the geometric multiplicity of d (E u(d)) algebraic multiplicity.</li> <li>If A is symmetric than (the eigenvalue) are real there expension of the subspace have there expension of the powerties are those with multiplicity in there expensions are those with multiplicity in there expensions are there is a symmetric then (the eigenvalues are real) there expensions distance characterization of the eigenvalues. As a de a set A multiplicity repeating then has a max multiplicity are subspace rank(s) = i</li> </ul>			
solving the polynomial equation det $(A - \lambda I) = 0$ If A hos $r \leq n$ engenvalues $d_{i_1, \ldots, i_r}$ then $i = n$ $(\lambda_j)$ $\lambda = det (A - \lambda I) = \prod_{j=1}^{r} (\lambda_j - \lambda_j)$ positive ologebraic $ologebraic nu(h) = n(\lambda_j)\lambda = det (A - \lambda I) = \prod_{j=1}^{r} (\lambda_j - \lambda_j)ologebraic nu(h) = n(\lambda_j)i = $	•	We can find all the eigenvalues of A by	
det $(A - II) = 0$ If A has $r \leq n$ eigenvalue $\lambda_{i_1, \dots, i_r}$ then $\lambda \mapsto det (A - \lambda I) = T(\lambda_j - \lambda)$ is objective mithicuts of mithicuts of $\lambda_j$ If $\lambda$ is on eigenvalue of A, then the dimension of the subspace $null (A - \lambda I_n)$ is the geometric multiplicity of $\lambda (\leq u(\lambda))$ $\lambda_j \in braic mithicuts of \lambda_j\lambda_j = max min \mu T A \lambda\lambda_n = max min \mu T A \lambda\lambda_n = max min \mu T A \lambda\lambda_n = max \mu = S\mu = K(S) = f$		solving The polynomial equation	•
If A has rear eigenvalue $\lambda_1,, \lambda_r$ then $\lambda_r$ $\mu(\lambda_j)$ $\lambda_r$ $det (A - \lambda_I) = \prod_{j=1}^r (J_j - \lambda_j) \int_{J_j}^{J_j} pointing r$ olgebrac = olgebrac = olgebr		det(A-JI)=0	
If A has $r \leq n$ eigenvalues $\lambda_i, \dots, \lambda_r$ Then $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) = T(A) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) + \mu(\lambda_i) + \mu(\lambda_i) + \mu(\lambda_i)$ $\lambda \mapsto det (A - \lambda I) + \mu(\lambda_i) + \mu(\lambda$			•
Then $\lambda \mapsto \det (A - \lambda I) = \prod_{j=1}^{r} (\lambda_j - \lambda_j)$ $j = \prod_{j=1}^{r$		lf A has rise nuclues di,, dr	•
A is on eigenvalue of A, then the dimension of the subspace null (A - A In)     is the geometric multiplicity of A (S = u(A))     is the geometric multiplicity of A (S = u(A))     is the geometric multiplicity of A (S = u(A))     is the geometric then (the eigenvalue are indiplicity     is symmetric then (the eigenvalue are indiplicity     there exists a variational characterization of     it eigenvalues is A = d = = = An possibly     repeating     Then     A = max     min xTAn     S = R <sup>n</sup> n = S     lineor     integer     is pose in multiplicity		theor r m(dj)	
objectives of numplicity of Numplicity of dis if $d$ is on eigenvalue of $A$ , then the dimension of the subspace null $(A - dI_n)$ is the geometric multiplicity of $d$ $(\leq u(d))$ eigebraic simple eigenvalues are those with multiplicity if $A$ is symmetric then (the eigenvalue are real) there exists a variational characterization of the eigenvalues: $d_i \geq d_2 \geq \ldots \geq \lambda_n$ possibly repeating Then $\lambda_n = \max(n = S = 1)$ have $\max(s) = \hat{i}$		$\lambda \mapsto \operatorname{alet} (A - \lambda I) = \Pi (\lambda j - \lambda) \int \operatorname{positive}_{j=1}^{j=1} (\lambda j - \lambda) \int \operatorname{positive}_{j=1} (\lambda$	eſ
If $\lambda$ is on eigenvalue of $A$ , then the dimension of the subspace null $(A - \lambda T_n)$ is the geometric multiplicity of $\lambda$ $(\leq u(\lambda))$ algebraic simple eigenvalues are those with multiplicity i if $A$ is symmetric then (the eigenvalue are real) there exists a variational characterization of there eigenvalues $\lambda_i \geq d_2 \geq \ldots \geq \lambda_n$ possibly repeating Then $\lambda_n = mark$ min $xTAn$ $\lambda_n = mark$ min $xTAn$ $\lambda_n = mark$ min $xTAn$ halk = nark(s) = i	· · · · · ·	oulgebraic some mutiplicity of	•
$f_{1} = \frac{1}{2} = \frac{1}{2$			
dimension of the subspace null $(A - A \perp_n)$ is the geometric multiplicity of $d (\leq u(d))$ elgebraic simple eigenvalues are those with multiplicity i if $A$ is symmetric then (the eigenvalues are real) there exists a variational characterization of there exists a variational characterization of the eigenvalues: $\lambda_i \geq d_2 \geq \ldots \geq \lambda_n$ possibly repeating Then $\lambda_n = max$ min $x \top A x$ $\lambda_n = max$ min $x \top A x$ $\lambda_n = max$ $m \in S$ ki R = i ki R = i		it is an eigenvalue of it, then the	
is the geometric multiplicity of $d (\leq u(d))$ elgebraic numple eigenvalues are those with multiplicity i if $A$ is symmetric then (the eigenvalues are real) there exists a variational characterization of the eigenvalues: $\lambda_i \geq d_2 \geq \ldots \geq \lambda_n$ possibly repeating Then $\lambda_i = \max(min  x \in S)$ $\lim_{k \neq i \in S} \sum_{k \in S}$		dimension of the subspace null (A-din)	
• Simple eigenvalues are those with multiplicity i • If A is symmetric then (the eigenvalues are real) there exists a variational characterization of the eigenvalues. $\lambda_i \ge d_2 \ge \dots \ge \lambda_n$ possibly repeating Then $\lambda_n = \max $		is the geometric multiplicity of d ( ≤ u(d))	
• If A is symmetric then (the eigenvalue are real) there exists a variational characterization of the eigenvalues: $\lambda_i \ge \lambda_2 \ge \ldots \ge \lambda_n$ possibly repeating Then $\lambda_i = \max \min \pi TAn$ $\lambda_i = \max \max \min \pi TAn$ $\lim_{s \to s \in \mathbb{R}^n} n \in S$ $\lim_{s \to s \in \mathbb{R}^n} \max(s) = i$	· · · · •	Simple eigenvalues are those with multiplicity i	y
there exists $\geq$ variational characterization of it eigenvalues: $\lambda_i \geq \lambda_2 \geq \ldots \geq \lambda_n$ possibly repeating Then $\lambda_i \equiv \max$ min $\pi T A \pi$ $\lambda_i \equiv \max S \subseteq R^n$ $\pi \in S$ $\lim_{s \downarrow b \ space} \operatorname{rank}(S) = 1$	· · · · ·	17 A is symmetric then (the eigenvolve) are real	) }
then Then $\lambda_{1} = max$ min $\pi TAn$ $\lambda_{n} = max$ min $\pi TAn$ $\lambda_{n} = max$ min $\pi TAn$ $\lambda_{n} = max$ $\pi \in S$ $\lim_{s \to b} space$ $rank(S) = 1$		there exists a variational characterization of	•
Then $k_{1} = max$ min $xTAn$ $k_{2} = max$ $x \in S$ $k_{2}R^{n}$ $n \in S$ $k_{2}R^{n} = n$ $k_{2}R = n$ $k_{2}R = n$		$d_1 \ge d_2 \ge \ldots \ge \lambda_n$ possibly repeating	•
$\lambda i = max$ min $\pi TA \pi$ linear $S \subseteq IR^n$ $\pi \in S$ $HZK \equiv I$ SUBSPORE $rank(S) = 1$		Then	•
linear $S \subseteq \mathbb{R}^n$ $n \in S$ subspace rank $(S) = \hat{i}$		X Max min xTAx	•
kineor subspace rank $(S) = \hat{1}$		$S \subseteq \mathbb{R}^n$ $n \in S$	
	line	spore rank $(S) = 1$	0
		2	•
lineor TSR NET	lin	$eos = T \leq R$	
subspace $\neg$ rank $(\tau) = n - n \tau $ $  _{\tau}   _{\tau}$	21	$nspose$ rank $(\tau) = n - n \tau$ $  _{\tau}   _{\tau}$	
			) [