Basic identities of matrix/vector ops	<u>j</u> j	Vector norms (beyond euclidean)	Determinant of square-diagonals =>	If all else fails, try to find row/column with MOST zeros	If associated to same eigenvalue <u>\(\)</u> then eigenspace	$ \sigma_1,, \sigma_p $ are singular values of \underline{A}].	Variance (Bessel's correction) of α ₁ ,, α _m is
$(A+B)^T = A^T + B^T (AB)^T = B^T A^T (A^{-1})^T = (A^T)^{-1} $	Notice: $Q_j c_j = \sum_{i=1}^{j} (\mathbf{q}_i \cdot \mathbf{a}_{j+1}) \mathbf{q}_i = \sum_{i=1}^{j} \operatorname{proj}_{\mathbf{q}_i} (\mathbf{a}_{j+1})$, so	vector norms are such that: $ x = 0 \iff x = 0$,	$\left \begin{array}{c} \operatorname{diag}(a_1,, a_n) = \prod_i a_i \\ \operatorname{triangular matrices}) \end{array} \right $	Perform minimal EROs/ECOs to get that row/column to be all-but-one zeros	E_{λ} has spanning-set $\{x_{\lambda_i}, \dots\}$	(Positive) singular values are (positive) square-roots	$Var_{\mathbf{w}} = \frac{1}{m-1} \sum_{j} \alpha_{j}^{2} = \frac{1}{m-1} \mathbf{w}^{T} \left(\sum_{j} \mathbf{r}_{j}^{T} \mathbf{r}_{j} \right) \mathbf{w}$
$(AB)^{-1} = B^{-1}A^{-1}$	rewrite as	$\frac{ \lambda x = \lambda x }{ x + y \le x + y }$		Don't forget to keep track of sign-flipping &	$x_1,, x_n$ are linearly independent \Rightarrow apply Gram-Schmidt $q_{\lambda_i}, \leftarrow x_{\lambda_i},$	of eigenvalues of $\underline{AA^T}$ or $\underline{A^TA}$ i.e. $\sigma_1^2,, \sigma_D^2$ are eigenvalues of $\underline{AA^T}$ or $\underline{A^TA}$	$= \frac{1}{m-1} \mathbf{w}^T \mathbf{A}^T \mathbf{A} \mathbf{w}$
For $\underline{A \in \mathbb{R}^{m \times n}} A_{ij}$ is the i -th ROW then j -th COLUMN	j j	$ \mathbf{r}_p $ norms: $ \mathbf{x} _p = (\sum_{i=1}^n \mathbf{x}_i ^p)^{1/p}$	The (column) rank of AJ is number of linearly		Then $\{\mathbf{q}_{\lambda_i}, \dots\}$ is orthonormal basis (ONB) of E_{λ_i}		First (principal) axis defined =>
$(A^{T})_{ij} = A_{ji} (AB)_{ij} = A_{i\star} \cdot B_{\star j} = \sum_{i} A_{ik} B_{kj} $	$\mathbf{u}_{j+1} = \mathbf{a}_{j+1} - \sum_{i=1} (\mathbf{q}_i \cdot \mathbf{a}_{j+1}) \mathbf{q}_i = \mathbf{a}_{j+1} - \sum_{i=1} \operatorname{proj}_{\mathbf{q}_i} (\mathbf{a}_{j+1})$	$p = 1$: $\ \mathbf{x}\ _1 = \sum_{i=1}^n \mathbf{x}_i $	independent columns, i.e. rk(A) I.e. its the number of pivots in row-echelon-form	notice all-but-one minor matrix determinants go to	$Q = \langle \mathbf{q}_1,, \mathbf{q}_n \rangle$ is an ONB of $\mathbb{R}^n \Longrightarrow \mathbb{Q} = [\mathbf{q}_1 \mathbf{q}_n]$ is	Let r = rk(A), then number of strictly positive singular	$\mathbf{w}_{(1)} = \arg \max_{\ \mathbf{w}\ = 1} \mathbf{w}^T \mathbf{A}^T \mathbf{A} \mathbf{w}$
R	$[a_1, \dots, a_n \in \mathbb{R}^m]$ $[m \ge n]$	$\frac{p=2}{2} \ \mathbf{x}\ _2 = \sqrt{\sum_{i=1}^n \mathbf{x}_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}$	I.e. its the dimension of the column-space	Representing EROs/ECOs as transfor-	orthogonal matrix i.e. $Q^{-1} = Q^{T}$	values is r	= arg max $\ \mathbf{w}\ = 1 (m-1) \text{Var}_{\mathbf{w}} = \mathbf{v}_1$ i.e. $\mathbf{w}_{(1)}$ the direction that maximizes variance $\text{Var}_{\mathbf{w}}$.
$(Ax)_i = A_{i*} \cdot x = \sum_j A_{ij} x_j \left[x^T y = y^T x = x \cdot y = \sum_i x_i y_i \right]$	<u>n</u> j	$\frac{p = \infty}{\ \mathbf{x}\ _{\infty} = \lim_{p \to \infty} \ \mathbf{x}\ _{p} = \max_{1 \le i \le n} \mathbf{x}_{i} }$	rk(A) = dim(C(A)) I.e. its the dimension of the image-space	mation matrices	$ \mathbf{q}_1, \dots, \mathbf{q}_n $ are still eigenvectors of $\underline{A} = A = QDQ^T$	ti.e. $\sigma_1 \ge \cdots \ge \sigma_r > 0$ and $\sigma_{r+1} = \cdots = \sigma_p = 0$	i.e. maximizes variance of projections on line Rw (1)
$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_i \sum_j \mathbf{A}_{ij} \mathbf{x}_i \mathbf{x}_j \mathbf{x}_i$ $\mathbf{x} \mathbf{e}_{R}^T = [0 \dots \mathbf{x} \dots 0]$	$(q_1,,q_n) \in \mathbb{R}^m \text{for } U_n \subset \mathbb{R}^m $	Any two norms in Rn are equivalent, meaning there	$rk(A) = dim(im(f_A))$ of linear map $f_A(x) = Ax$	For A ∈ R ^{m×n} suppose a sequence of:	(spectral decomposition)	$+A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^I$	σ ₁ u ₁ ,,σ _r u _r (columns of <u>US</u>) are principal
$\mathbf{e}_{k}\mathbf{x}^{T} = [0^{T}; \dots; \mathbf{x}^{T}; \dots; 0^{T}]$	$j=1 \Rightarrow u_1 = a_1$ and $q_1 = \hat{u}_1$, i.e. start of iteration	exist $r>0$; $s>0$ such that: $\forall x \in \mathbb{R}^{n}$, $r\ x\ _{a} \le \ x\ _{b} \le s\ x\ _{a}$	The (row) rank of AJis number of linearly independent	$ EROs \text{ transform } \underline{A} \rightsquigarrow_{EROs} \underline{A'} \Rightarrow \text{ there is matrix } \underline{R} \text{s.t.}$ $ RA = A' $	A = QDQ ^T can be interpreted as scaling in direction of its eigenvectors:	SVD is similar to spectral decomposition, except it	components/scores of A
Scalar-multiplication + addition distributes over:	$ j=2 \Rightarrow \frac{\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{q}_1 \cdot \mathbf{a}_2)\mathbf{q}_1}{\mathbf{q}_2 = \mathbf{u}_2} $ and $ \mathbf{q}_2 = \mathbf{u}_2 $ etc Linear independence guarantees that $\mathbf{a}_{j+1} \notin U_j$	x _∞ ≤ x ₁	The row/column ranks are always the same, hence	ECOs transform A → ECOs A' => there is matrix CJs.t.	Perform a succession of reflections/planar rotations	always exists	Recall: $A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ with $\sigma_1 \ge \cdots \ge \sigma_r > 0$, so that
column-blocks => $ A + B = \lambda [A_1 A_C] + [B_1 B_C] = [\lambda A_1 + B_1 \lambda A_C + B_C] $	For exams: compute $\mathbf{u}_{j+1} = \mathbf{a}_{j+1} - Q_j \mathbf{c}_j$	Equivalence of ℓ_1, ℓ_2 and $\ell_\infty \Rightarrow \ \mathbf{x}\ _2 \le \sqrt{n} \ \mathbf{x}\ _\infty$	$rk(A) = dim(C(A)) = dim(R(A)) = dim(C(A^T)) = rk(A^T)$	AC = A'	to change coordinate-system -Apply scaling by λ _i to each dimension q _i	If $\underline{n \le m}$ then work with $\underline{A^T A \in \mathbb{R}^{n \times n}}$.	relates principal axes and principal components
row-blocks ⇒	Gather $Q_j = [\mathbf{q}_1 \mid \mid \mathbf{q}_j] \in \mathbb{R}^{m \times j}$	$\ \mathbf{x}\ _1 \le \sqrt{n} \ \mathbf{x}\ _2$	A is full-rank iff rk(A) = min(m, n), i.e. its as linearly independent as possible	Both transform A → EROS+ECOS A' => there are matrices R, C s.t. RAC = A'	-Undo those reflections/planar rotations	Obtain eigenvalues $\sigma_1^2 \ge \cdots \ge \sigma_n^2 \ge 0$ of $A^T A$ Obtain orthonormal eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ of	Data compression: If o ₁ ≫ o ₂ then compress Alby projecting in direction of principal component =>
$\underbrace{\lambda A + B = \lambda [A_1;; A_r] + [B_1;; B_r] = [\lambda A_1 + B_1;; \lambda A_r + B_r]}_{}$	Compute $\mathbf{c}_j = [\mathbf{q}_1 \cdot \mathbf{a}_{j+1}, \dots, \mathbf{q}_j \cdot \mathbf{a}_{j+1}]^T \in \mathbb{R}^j$	Induce metric $d(x,y) = y-x $ has additional properties:		matrices k, C s.t. KAC = A	Extension to C ⁿ	$\underline{A^T A}$ (apply normalization e.g. Gram-Schmidt !!!! to	$A \approx \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T$
Matrix-multiplication distributes over: column-blocks \Rightarrow AB = A[B ₁ B _D] = [AB ₁ AB _D]	Compute $Q_j c_j \in \mathbb{R}^m$ and subtract from a_{j+1}	Translation invariance: $d(x+w, y+w) = d(x, y)$	Two matrices $\mathbf{A}, \tilde{\mathbf{A}} \in \mathbb{R}^{m \times n}$ are equivalent if there exist	FORWARD: to compute these transformation	Standard inner product: $(x, y) = x^{\dagger} y = \sum_{i} \overline{x_{i}} y_{i}$	eigenspaces E _G ,)	
	Duran autient dat aus durat 6 aranga	Scaling: $d(\lambda x, \lambda y) = \lambda d(x, y)$ Matrix norms	two invertible matrices $\underline{P} \in \mathbb{R}^{m \times m}$ and $\underline{Q} \in \mathbb{R}^{n \times n}$ such that $\underline{A} = \underline{P} \underline{A} \underline{O}^{-1}$	matrices: Start with [I _m A I _n] i.e. A Jand identity matrices	Conjugate-symmetric: $(x, y) = \overline{(y, x)}$	$V = [\mathbf{v}_1 \mid \dots \mid \mathbf{v}_n] \in \mathbb{R}^{n \times n}$ is orthogonal so $V^T = V^{-1}$	Cholesky Decomposition Consider positive (semi-)definite $A \in \mathbb{R}^{n \times n}$
outer-product sum =>	$x^{T}y = y^{T}x = x \cdot y = \sum_{i} x_{i}y_{i} x \cdot y = a b \cos x\hat{y}$	Matrix norms are such that: $ A = 0 \iff A = 0$	Two matrices $\mathbf{A}, \tilde{\mathbf{A}} \in \mathbb{R}^{n \times n}$ are similar if there exists an	For every ERO on Al, do the same to LHS (i.e. I _m)	Standard (induced) norm: $ x = \sqrt{\langle x, y \rangle} = \sqrt{x^{\dagger} y}$	$r = rk(A) = no.$ of strictly +ve σ_i	Cholesky Decomposition is A = LLT where L is
$AB = [A_1 A_p][B_1;; B_p] = \sum_{i=1}^{p} A_i B_i$	i $x \cdot y = y \cdot x \mid x \cdot (y + z) = x \cdot y + x \cdot z \mid \alpha x \cdot y = \alpha(x \cdot y)$	$ \lambda A = \lambda A $, $ A+B \le A + B $	invertible matrix $P \in \mathbb{R}^{n \times n}$ such that $A = P\tilde{A}P^{-1}$	For every ECO on <u>A</u> J do the same to RHS (i.e. $\overline{I_n}$) Once done, you should get $[I_m \mid A \mid I_n] \rightsquigarrow [R \mid A \mid C]$	We can diagonalise real matrices in CJwhich lets us diagonalise more matrices than before	Let $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ then $\underline{\mathbf{u}_1,, \mathbf{u}_r \in \mathbb{R}^m}$ are orthonormal	lower-triangular For positive semi-definite => always exists, but
e.g. for $A = [a_1 \mid \mid a_n]$ $B = [b_1;; b_n] \Longrightarrow AB = \sum_i a_i b_i$	$x \cdot x = x ^2 = 0 \iff x = 0$	Matrices matrix norms are vector norms, all results apply	Similar matrices are equivalent, with Q = P A is diagonalisable iff A is similar to some diagonal	with RAC = A'	Least Square Method	(therefore linearly independent) The orthogonal compliment of span{u ₁ ,,u _r } =>	non-unique
	for $\underline{x \neq 0}$, we have $\underline{x \cdot y = x \cdot z} \Longrightarrow x \cdot (y - z) = 0$	Sub-multiplicative matrix norm (assumed by default)	matrix D	If the sequences of EROs and ECOs were $R_1,, R_{\lambda}$ and	If we are solving Ax = b and b ∉ C(A) i.e. no solution,	$span\{\mathbf{u}_1,, \mathbf{u}_r\}^{\perp} = span\{\mathbf{u}_{r+1},, \mathbf{u}_m\}$	For positive-definite => always uniquely exists s.t. diagonals of L are positive
	$ x \cdot y \le x y $ (Cauchy-Schwartz inequality) $ u+v ^2 + u-v ^2 = 2 u ^2 + 2 v ^2$ (parallelogram law)	is also such that AB ≤ A B	Properties of determinants	$C_1,,C_{\mu}$ respectively	then Least Square Method is: Finding xjwhich minimizes Ax-b ₂	*Solve for unit-vector u _{r+1} s.t. it is orthogonal to	
idempotent)	$ u+v \le u + v $ (triangle inequality)	Common matrix norms, for some $\underline{\mathbf{A}} \in \mathbb{R}^{m \times n}$: $\ \mathbf{A}\ _{1} = \max_{j} \ \mathbf{A}_{*j}\ _{1}$	Consider $\underline{A \in \mathbb{R}^{n \times n}}$, then $A_{ij}' \in \mathbb{R}^{(n-1) \times (n-1)}$ the	$R = R_{\lambda} \cdots R_{1}$ and $C = C_{1} \cdots C_{\mu}$ so	Recall for $A \in \mathbb{R}^{m \times n}$ we have unique decomposition	*Then solve for unit-vector \mathbf{u}_{r+2} s.t. it is orthogonal	Finding a Cholesky Decomposition:
projection matrix	$u \perp v \iff u+v ^2 = u ^2 + v ^2$ (pythagorean	$\ \mathbf{A}\ _2 = \sigma_1(\mathbf{A})$ i.e. largest singular value of \mathbf{A}	[(i,j)]-minor matrix of A] obtained by deleting i-th row and j-th column from A	$(R_{\lambda} \cdots R_1)A(C_1 \cdots \overline{C_{\mu}}) = A'$	for any $\mathbf{b} \in \mathbb{R}^m$ $\mathbf{b} = \mathbf{b}_i + \mathbf{b}_k$	to u ₁ ,, u _{r+1}	Compute <u>LL^T</u> and solve <u>A=LL^T</u> by matching terms For square roots always pick positive
It is called an orthogonal projection matrix if	theorem) $\ c\ ^2 = \ a\ ^2 + \ b\ ^2 - 2\ a\ \ b\ \cos b\hat{a} (law of cosines)$	(square-root of largest eigenvalue of A ^T A or AA ^T	Then we define determinant of A ₃ i.e. det(A) = A , as	$R^{-1} = R_1^{-1} \cdots R_{\lambda}^{-1}$ and $C^{-1} = C_{\mu}^{-1} \cdots C_{1}^{-1}$, where	where $\mathbf{b}_i \in C(A)$ and $\mathbf{b}_k \in \ker(A^T)$	*And so on $U = [\mathbf{u}_1 \dots \mathbf{u}_m] \in \mathbb{R}^{m \times m} \text{ is orthogonal so } \underline{U}^T = \underline{U}^{-1}$	If there is exact solution then positive-definite
P = P = P (conjugate-transpose)	Transformation matrix & linear maps	$\ \mathbf{A}\ _{\infty} = \max_{i} \ \mathbf{A}_{i\star}\ _{1}$, note that $\ \mathbf{A}\ _{1} = \ \mathbf{A}^{T}\ _{\infty}$	$\det(A) = \sum_{i=1}^{n} (-1)^{i+k} A_{ik} \det(A_{ik}')$ i.e. expansion along	R_i^{-1}, C_j^{-1} are inverse EROs/ECOs respectively	$\ A\mathbf{x} - \mathbf{b}\ _2$ is minimized $\iff \ A\mathbf{x} - \mathbf{b}_i\ _2 = 0 \iff A\mathbf{x} = \mathbf{b}_i$		If there are free variables at the end, then positive semi-definite
Because π: V → V lis a linear map , its image space	For linear map $f: \mathbb{R}^n \to \mathbb{R}^m$, ordered bases	Frobenius norm: $\ \mathbf{A}\ _F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{A}_{ij} ^2}$	R=1		$A^T A = A^T b$ is the normal equation which gives	$S = \text{diag}_{m \times n}(\sigma_1,, \sigma_n)$ AND DONE!!! If $m < n$ then let $B = A^T$	i.e. the decomposition is a solution-set
$U = im(\pi)$ and null space $W = ker(\pi)$ are subspaces of V	$(\mathbf{b}_1, \dots, \mathbf{b}_n) \in \mathbb{R}^m$ and $(\mathbf{c}_1, \dots, \mathbf{c}_m) \in \mathbb{R}^m$	Frobenius norm: $\ \mathbf{A}\ _F = \sqrt{\sum_{i=1}^{r} \sum_{j=1}^{r} \mathbf{A}_{ij} }$	ijth row *(for any i)	BACKWARD: once $R_1,,R_{\lambda}$ and $C_1,,C_{\mu}$ for which	solution to least square problem:	apply above method to $B \implies B = A^T = USV^T$	parameterized on free variables
	$A = \mathbf{F}_{CB} \in \mathbb{R}^{m \times n}$ is the transformation-matrix of f w.r.t to bases B and C	A matrix norm $\ \cdot\ $ on $\mathbb{R}^{m \times n}$ is consistent with the	$\det(A) = \sum_{i=1}^{n} (-1)^{R+j} A_{kj} \det(A_{kj}')$, i.e. expansion along	$RAC = A'$ are known , starting with $[I_m \mid A \mid I_n]$	$\ \mathbf{A}\mathbf{x} - \mathbf{b}\ _2$ is minimized $\iff \mathbf{A}\mathbf{x} = \mathbf{b}_i \iff \mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T \mathbf{b}$	$A = B^T = VS^TU^T$	e.g. 1 1 1 = LL ^T where L = 1 0 0 , c ∈ [0,1]
$W = im(\pi^*) = ker(\pi)$ and $U = ker(\pi^*) = im(\pi)$, i.e. they	$f(\mathbf{b}_j) = \sum_{i=1}^m A_{ij} \mathbf{c}_i \rightarrow \text{ each } \mathbf{b}_j \text{ basis gets mapped to a}$	vector norms $\ \cdot \ _a$ on \mathbb{R}^n and $\ \cdot \ _b$ on \mathbb{R}^m if	k=1 j th column (for any j)	For $\underline{i=1 \rightarrow \lambda}$ perform $\underline{R_i}$ on \underline{A} , perform $\underline{R_{\lambda-j+1}}^{-1}$ on LHS	Linear Regression	Tricks: Computing orthonormal	[1 1 2] [1 c √1-c ²]
swapped ∤πJis a projection along WJonto UJ	linear combination of $\sum_i a_i c_i$ bases	for all $\underline{\mathbf{A}} \in \mathbb{R}^{m \times n}$ and $\underline{\mathbf{x}} \in \mathbb{R}^n \Longrightarrow \ \underline{\mathbf{A}}\mathbf{x}\ _b \le \ \underline{\mathbf{A}}\ \ \mathbf{x}\ _a$ If $a = b$, $\ \cdot\ $ is compatible with $\ \cdot\ _a$	When det(A) = 0 we call AJa singular matrix	(i.e. I _m)	Let $y = f(t) = \sum_{j=1}^{n} s_j f_j(t)$ be a mathematical model,	vector-set extensions You have orthonormal vectors $\mathbf{u}_1,,\mathbf{u}_r \in \mathbb{R}^m$ ⇒ need	If A = LL ^T you can use forward/backward substitution
-π* is a projection along U onto W	If f^{-1} exists (i.e. its bijective and $m = n$) then	Frobenius norm is consistent with \(\ell_2\) norm \(\Rightarrow\)	Common determinants For n = 1, det(A) = A ₁₁	For $\underline{j=1 \rightarrow \mu}$ perform $C_{\underline{j}}$ on \underline{A} , perform $C_{\underline{\mu-j+1}}^{-1}$ on RHS (i.e. I_n)	where f_j are basis functions and s_j are parameters Let $(t_i, y_i) \mid 1 \le i \le m, m \gg n$ be a set of observations,	to extend to orthonormal vectors $\mathbf{u}_1,, \mathbf{u}_m \in \mathbb{R}^m$	to solve equations
	$(\mathbf{F}_{CB})^{-1} = \mathbf{F}^{-1}_{BC}$ (where \mathbf{F}^{-1}_{BC} is the transformation-matrix of f^{-1})	Av ₂ ≤ A _F v ₂	For n = 21, det(A) = A ₁₁ A ₂₂ - A ₁₂ A ₂₁	You should get $[I_m \mid A \mid I_n] \rightarrow [R^{-1} \mid A' \mid C^{-1}]$ with	and $\mathbf{t}, \mathbf{y} \in \mathbb{R}^m$ are vectors representing those	Special case \Rightarrow two 3D vectors \Rightarrow use cross-product \Rightarrow $a \times b \perp a, b \mid$	For $\underline{Ax = b} \Longrightarrow \text{let } \underline{y = L^T x}$
vector <u>x ∈ V</u> can be uniquely written as <u>x = u + w</u>	transformation-matrix or jy	For a vector norm $\ \cdot\ $ on \mathbb{R}^n , the subordinate matrix norm $\ \cdot\ $ on $\mathbb{R}^{m \times n}$ is	det(I _n)=1	A=R ⁻¹ A'C ⁻¹	observations		Solve $Ly = b$ by forward substitution to find y Solve $L^T x = y$ by backward substitution to find x
	The transformation matrix of the identity map is called	$\ A\ = \max\{\ Ax\ : x \in \mathbb{R}^n, \ x\ = 1\}$	Multi-linearity in columns/rows: if $A = [a_1 a_i a_n] = [a_1 \lambda x_i + \mu y_i a_n] \text{ [then]}$		$f_j(\mathbf{t}) = [f_j(\mathbf{t}_1), \dots, f_j(\mathbf{t}_m)]^T$ is transformed vector $A = [f_1(\mathbf{t})] \dots f_n(\mathbf{t}) \in \mathbb{R}^{m \times n}$ is a matrix of columns	Extension via standard basis $I_m = [e_1 e_m]$ using	[I ₁₁ 0 0]
" <u> </u>	change-in-basis matrix The identity matrix I _m represents id _R m w.r.t. the	$= \max \left\{ \frac{\ \mathbf{A}\mathbf{x}\ }{\ \mathbf{x}\ } : \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0 \right\}$	$\det(A) = \lambda \det\left(\begin{bmatrix} a_1 \\ \dots \\ x_i \end{bmatrix} \dots \begin{bmatrix} a_n \end{bmatrix} \right)$	You can mix-and-match the forward/backward modes i.e. inverse operations in inverse order for one, and	$\mathbf{z} = [s_1, \dots, s_n]^T$ is vector of parameters	(tweaked) GS: Choose candidate vector: just work through	For <u>n = 3</u>]=> L = l ₂₁ l ₂₂ 0
i.e. the image and kernel of <u>π</u> are orthogonal	standard basis $E_m = \langle e_1,, e_m \rangle \Rightarrow i.e. I_m = I_{EE}$	$= \max\{\ \mathbf{A}\mathbf{x}\ : \mathbf{x} \in \mathbb{R}^n, \ \mathbf{x}\ \le 1\}$	+µ det ([a ₁ y _i a _n])	operations in normal order for the other	Then we get equation Az = y => minimizing Az - y ₂	e ₁ ,, e _m sequentially starting from e ₁ \Rightarrow denote	[l31 l32 l33]
subspaces infact they are eachother's orthogonal compliments,	If $B = \langle \mathbf{b}_1, \dots, \mathbf{b}_m \rangle$ is a basis of \mathbb{R}^m , then	Vector norms are compatible with their subordinate	And the exact same linearity property for rows	e.g. you can do $[I_m \mid A \mid I_n] \rightsquigarrow [R^{-1} \mid A' \mid C]$ to get $AC = R^{-1}A' \implies U$ useful for LU factorization	is the solution to Linear Regression	the current candidate e_k Orthogonalize: Starting from $j = r$ going to $j = m$ with	$LL^T = \begin{bmatrix} l_{11}^c & l_{11}l_{21} & l_{11}l_{31} \\ l_{11}l_{21} & l_{21}^2 * l_{22}^2 & l_{21}l_{31} * l_{22}l_{32} \end{bmatrix}$
i.e. $U^{\perp} = W, W^{\perp} = U$ (because finite-dimensional	to E	matrix norms For $p = 1, 2, \infty$ matrix norm $\ \cdot\ _p$ is subordinate to	M = M	Eigen-values/vectors	So applying LSM to Az = y is precisely what Linear Regression is	each iteration => with current orthonormal vectors	LL' = 11121 121 +122 12131 +122132 11131 12131 +122132 131 +132 +133
vectorspaces) so we have $\pi(x) \cdot y = \pi(x) \cdot \pi(y) = x \cdot \pi(y)$	$I_{BE} = (I_{EB})^{-1}$, so $\Rightarrow F_{CB} = I_{CE}F_{EE}I_{EB}$	the vector norm $\ \cdot\ _D$ (and thus compatible with)	Alternating: if any two columns of Alare equal (or any	Consider $A \in \mathbb{R}^{n \times n}$, non-zero $\mathbf{x} \in \mathbb{C}^n$ is an eigenvector	-We can use normal equations for this =>	u ₁ ,,u _j +Compute	Forward/backward substitution
or equivalently, $\pi(x) \cdot (y - \pi(y)) = (x - \pi(x)) \cdot \pi(y) = 0$		Properties of matrices	two rows of A are equal), then A = 0 (its singular)	with eigenvalue $\lambda \in \mathbb{C}$ for A if $Ax = \lambda x$	$ Az-y _2$ is minimized $\iff A^TAz=A^Ty$	111 : : : : :	Forward substitution: for lower-triangular
By Cauchy–Schwarz inequality we have ∥π(x)∥ ≤ ∥x∥	Dot-product uniquely determines a vector w.r.t. to basis	Consider <u>A</u> ∈ R ^{m×n}	Immediately from this (and multi-linearity) => if columns (or rows) are linearly-dependent (some are	If $Ax = \lambda x$ then $A(kx) = \lambda(kx)$ for $k \neq 0$, i.e. kx is also an eigenvector	Solution to normal equations unique iff AJis full-rank, i.e. it has linearly-independent columns	$\mathbf{w}_{j+1} = \mathbf{e}_k - \sum_{i=1}^{J} (\mathbf{e}_k \cdot \mathbf{u}_i) \mathbf{u}_i = \mathbf{e}_k - \sum_{i=1}^{J} (\mathbf{u}_i)_k \mathbf{u}_i$ $= \mathbf{e}_k - U_i \mathbf{c}_i$	$\begin{bmatrix} \ell_{1,1} & 0 \\ L = \begin{bmatrix} \vdots & \ddots & \end{bmatrix} \end{bmatrix}$
The orthogonal projection onto the line containing	If $a_i = x \cdot b_i$; $x = \sum_i a_i b_i$, we call a_i the coordinate-vector of x_i w.r.t. to a_i	If <u>Ax = x</u> for all <u>x</u> then <u>A = I</u> For square <u>A</u> , the trace of <u>A</u> is the sum if its diagonals ,	linear combinations of others) then A = 0	AJhas at most nJdistinct eigenvalues	,	Where $U_i = [\mathbf{u}_1 \dots \mathbf{u}_i]$ and $\mathbf{c}_i = [(\mathbf{u}_1)_k, \dots, (\mathbf{u}_i)_k]^T$	$\begin{bmatrix} L^{-} & \vdots & \ddots \\ \ell_{n,1} & \dots & \ell_{n,n} \end{bmatrix}$
vector \underline{u} jis $\underline{proj}_{\underline{u}} = \hat{u}\hat{u}^T$, i.e. $\underline{proj}_{\underline{u}}(v) = \frac{\underline{u} \cdot v}{\underline{u} \cdot \underline{u}} u$; $\hat{u} = \frac{\underline{u}}{\ \underline{u}\ }$	Rank-nullity theorem:	i.e. tr(A)	Stated in other terms \Rightarrow rk(A) < n \iff A = 0 <=> RREF(A) \neq I _n \iff A = 0 (reduced row-echelon-form)	·The set of all eigenvectors associated with eigenvalue λ] is called eigenspace Ε _λ of <u>A</u>]	Positive (semi-)definite matrices	NOTE: e.b. : 11: = (11:) b. li.e. k!-th component of 11: 1	For Lx = b], just solve the first row
A special case of $\pi(x) \cdot (y - \pi(y)) = 0$ is $u \cdot (v - \text{proj}_{u} v) = 0$ since $\text{proj}_{u}(u) = u$	$\dim(\operatorname{im}(f)) + \dim(\ker(f)) = \operatorname{rk}(A) + \dim(\ker(A)) = n$	A] is symmetric iff $A = A^T$ A] is Hermitian, iff $A = A^{\dagger}$ i.e.	\Leftrightarrow $C(A) \neq \mathbb{R}^n \iff A = 0 (column-space)$	-E _λ = ker(A - λ/)	Consider symmetric $\underline{A \in \mathbb{R}^{n \times n}}$, i.e. $\underline{A = A^T}$ A is positive-definite $\underline{iff} x^T Ax > 0$ for all $x \neq 0$	-If $\mathbf{w}_{j+1} = 0$ then $\mathbf{e}_k \in \text{span}\{\mathbf{u}_1,, \mathbf{u}_j\}$ => discard	$\ell_{1,1} x_1 = b_1 \implies x_1 = \frac{b_1}{\ell_{1,1}}$ and substitute down
If $U \subseteq \mathbb{R}^n$ is a k -dimensional subspace with	full-rank	its equal to its conjugate-transpose	For more equivalence to the above, see invertible	The geometric multiplicity of $\underline{\Lambda}$ is $\dim(E_{\overline{\Lambda}}) = \dim(\ker(A - \lambda I))$	AJis positive-definite iff all its eigenvalues are strictly	w _{j+1} choose next candidate e _{k+1} try this step	Then salve the second row
orthonormal basis (ONB) $\langle \mathbf{u}_1,, \mathbf{u}_k \rangle \in \mathbb{R}^m$	Orthogonality concepts	AA ^T and A ^T A are symmetric (and positive semi-definite)	matrix theorem Interaction with EROs/ECOs:	The spectrum $Sp(A) = \{\lambda_1,, \lambda_n\}$ of A is the set of all	positive +A is positive-definite => all its diagonals are strictly	again Normalize: w _{j+1} ≠0 so compute unit vector	$\ell_{2,1} x_1 + \ell_{2,2} x_2 = b_2 \implies x_2 = \frac{b_2 - \ell_{2,1} x_1}{\ell_{2,2}}$ and
	$u \perp v \iff u \cdot v = 0$ i.e. u_j and v_j are orthogonal u_j and v_j are orthonormal iff $u \perp v$, $ u = 1 = v $	For real matrices, Hermitian/symmetric are	Swapping rows/columns flips the sign	eigenvalues of Al	positive	$ \mathbf{u}_{j+1} = \hat{\mathbf{w}}_{j+1} $	substitute down
Orthogonal projection onto U is $\pi_U = UU$	$A \in \mathbb{R}^{n \times n}$ is orthogonal iff $A^{-1} = A^T$	equivalent conditions	Scaling a row/column by <u>\lambda</u> ≠ 0] will scale the determinant by \lambda] (by multi-linearity)	The characteristic polynomial of AJis	A_{j} is positive-definite => $\max(A_{ij}, A_{jj}) > A_{ij} $	Repeat: keep repeating the above steps, now with	and so on until all x _i are solved
-Can be rewritten as $\pi_U(v) = \sum_i (\mathbf{u}_i \cdot v)\mathbf{u}_i$	Columns of A = [a ₁ a _n] are orthonormal basis	Every eigenvalue $\lambda_{\underline{i}}$ of Hermitian matrices is real geometric multiplicity of $\lambda_{\underline{i}}$ = geometric multiplicity	Remember to scale by λ^{-1} to maintain equality, i.e.	$\frac{P(\lambda) = A - \lambda I = \sum_{i=0}^{n} a_i \lambda^i}{a_0 = A \mid_{A_{n-1}} = (-1)^{n-1} \operatorname{tr}(A) \mid_{A_n} = (-1)^n}$	i.e. strictly larger coefficient on the diagonals AJis positive-definite => all upper-left submatrices are	new orthonormal vectors $\underline{\mathbf{u}_1,,\mathbf{u}_{j+1}}$	Backward substitution: for upper-triangular
If $\langle \mathbf{u}_1, \dots, \mathbf{u}_k \rangle$ is not orthonormal , then "normalizing	(ONB) $C = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle \in \mathbb{R}^n$ so $A = \mathbf{I}_{EC}$ is change-in-basis matrix	of λ _i	$\frac{\det(A) = \lambda^{-1} \det([a_1 \dots \lambda a_i \dots a_n])}{\det(A) = \lambda^{-1} \det([a_1 \dots \lambda a_i \dots a_n])}$ Invariant under addition of rows/columns	λ∈C is eigenvalue of A J iff λ is a root of P(λ)	also positive-definite	SVD Application: Principal Compo-	U = ·. :
factor" $(\mathbf{U}^T \mathbf{U})^{-1}$ is added $\Rightarrow \pi_U = \mathbf{U}(\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T$	Orthogonal transformations preserve	eigenvectors x_1, x_2 associated to distinct eigenvalues λ_1, λ_2 are orthogonal , i.e. $x_1 \perp x_2$	Invariant under addition of rows/columns Link to invertable matrices => $ A^{-1} = A ^{-1}$ which	The algebraic multiplicity of $\underline{\lambda}$] is the number of times it is repeated as root of $P(\lambda)$	Sylvester's criterion: Alis positive-definite iff all upper-left submatrices have strictly positive	nent Analysis (PCA) Assume $A_{uncentered} \in \mathbb{R}^{m \times n}$ represent <u>m</u> samples of	[0 u _{n,n}]
For line subspaces $U = \text{span}\{u\}$, we have $(\mathbf{U}^T \mathbf{U})^{-1} = (u^T u)^{-1} = 1/(u \cdot u) = 1/\ u\ $	lengths/angles/distances \Rightarrow $ Ax _2 = x _2$, $AxAy = xy$ Therefore can be seen as a succession of reflections	<u> </u>	means A is invertible ⇔ A ≠ 0 i.e. singular	1]s geometric multiplicity of \(\lambda \)	determinant	n -dimensional data (with $m \ge n$)	For $\underline{Ux = b}$ just solve the last row $u_{n,n} x_n = b_n \implies x_n = \frac{b_n}{u_{n,n}} \text{ and substitute up}$
Gram-Schmidt (GS) to gen. ONB from	and planar rotations	A jis triangular iff all entries above (lower-triangular) or	matrices are not invertible For block-matrices:	≤ algebraic multiplicity of λ	AJis positive semi-definite iff $x^T Ax \ge 0$ for all x	Data centering: subtract mean of each column from that column's elements	Then solve the second-to-last row
lin. ind. vectors	$\det(A) = 1$ or $\det(A) = -1$, and all eigenvalues of A jare	below (upper-triangular) the main diagonal are zero Determinant $\Rightarrow A = \prod_i a_{ii} \mid_i i.e.$ the product of	$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A) \det(B) = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix}$	Let $\lambda_1,, \lambda_n \in \mathbb{C}$ be (potentially non-distinct)	AJis positive semi-definite iff all its eigenvalues are	Let the resulting matrix be $\underline{A \in \mathbb{R}^{m \times n}}$, who's columns	$ u_{n-1} _{n-1} \times u_{n-1} _{n} \times u_{n-1$
Gram-Schmidt is iterative projection => we use	s.t. $ \lambda = 1$ $A \in \mathbb{R}^{m \times n}$ is semi-orthogonal iff $A^T A = I$ or $AA^T = I$	diagonal elements		eigenvalues of \underline{A}_j with $\underline{x_1,, x_n \in \mathbb{C}^n}$ their eigenvectors	non-negative -A_is positive semi-definite => all its diagonals are	have mean zero PCA is done on centered data-matrices like A:	$\Rightarrow x_{n-1} = \frac{b_{n-1} - u_{n-1,n-1} x_{n-1}}{u_{n-1,n}}$ and substitute up
current j dim subspace, to get next (j + 1) dim subspace	If <u>n > m</u> then all <u>m</u> rows are orthonormal vectors	A_Jis diagonal iff A _{ij} = 0, i ≠ j i.e. if all off-diagonal	$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(D - CA^{-1}B) $ if Alor D are	$tr(A) = \sum_{i} \lambda_{i}$ and $det(A) = \prod_{i} \lambda_{ij}$	non-negative	SVD exists i.e. $A = USV^T$ and $r = rk(A)$	and so on until all x _i are solved
Assume orthonormal basis (ONB) $(\mathbf{q_1}, \dots, \mathbf{q_j}) \in \mathbb{R}^m$	in in 2 in their att in cottainins are orthonormal vectors	entries are zero	= det(D) det(A - BD ⁻¹ C)	A jis diagonalisable iff there exist a basis of \mathbb{R}^n consisting of $\mathbf{x}_1, \dots, \mathbf{x}_n$	<u>A</u> Jis positive semi-definite => $\max(A_{ij}, A_{jj})$ ≥ $ A_{ij} $ i.e. no coefficient larger than on the diagonals	Let $A = [\mathbf{r}_1;; \mathbf{r}_m]$ be rows $\mathbf{r}_1,, \mathbf{r}_m \in \mathbb{R}^n$ \Rightarrow each row corresponds to a sample	" -
for j dim subspace $U_j \subset \mathbb{R}^m$	orthogonal subspaces	Written as $\operatorname{diag}_{m \times n}(\mathbf{a}) = \operatorname{diag}_{m \times n}(a_1, \dots, a_p), p = \min(m, n)$ where	invertible, respectively	consisting of $x_1,, x_n$ A) is diagonalisable iff $r_i = g_i$, where	AJis positive semi-definite => all upper-left	Let $A = [c_1 c_n]$ be columns $c_1,, c_n \in \mathbb{R}^m$ \Rightarrow each	Thin QR Decomposition w/ Gram-
Let $Q_j = [\mathbf{q}_1 \mid \mid \mathbf{q}_j] \in \mathbb{R}^{m \times j}$ be the matrix	Orthogonal compliment of $U \subset \mathbb{R}^n$ is the subspace	$\frac{\mathbf{a} = [a_1, \dots, a_p]^T \in \mathbb{R}^p \text{ diagonal entries of } \underline{\mathbf{A}}]}{\mathbf{a}}$	Sylvester's determinant theorem: $det(I_m + AB) = det(I_n + BA)$	r_i = geometric multiplicity of λ_i and g_i = geometric multiplicity of λ_i	submatrices are also positive semi-definite -Alis positive semi-definite => it has a Cholesky	column corresponds to one dimension of the data	Schmidt (GS) Consider full-rank $A = [a_1 a_n] \in \mathbb{R}^{m \times n} (m \ge n)$, i.e.
$P_j = Q_j Q_j^T$ is orthogonal projection onto U_j	$U^{\perp} = \{x \in \mathbb{R}^{n} \mid \forall y \in \mathbb{R}^{n} : x \perp y\}$ $= \{x \in \mathbb{R}^{n} \mid \forall y \in \mathbb{R}^{n} : x \le x + y \}$	For $x \in \mathbb{R}^n$ $Ax = \operatorname{diag}_{m \times n}(a_1, \dots, a_p)[x_1 \dots x_n]^T$ [if	Matrix determinant lemma:	Eigenvalues of A^k are $\lambda_1, \dots, \lambda_n$	Decomposition	Let $X_1,, X_n$ be random variables where each X_i corresponds to column c_i	$\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^m$ are linearly independent
$P_{\perp j} = I_m - Q_j Q_j^T$ is orthogonal projection onto	$\mathbb{R}^n = U \oplus U^{\perp}$ and $(U^{\perp})^{\perp} = U$	For $\underline{x \in \mathbb{R}^n}$ = $[a_1 x_1 \dots a_p x_p \ 0 \dots 0]^T \in \mathbb{R}^m$ (if	$\frac{\det (\mathbf{A} + \mathbf{u}\mathbf{v}^T) = (1 + \mathbf{v}^T \mathbf{A}^{-1}\mathbf{u}) \det(\mathbf{A})}{\det (\mathbf{A} + \mathbf{U}\mathbf{v}^T) = \det (\mathbf{I}_{\mathbf{M}} + \mathbf{v}^T \mathbf{A}^{-1}\mathbf{U}) \det(\mathbf{A})}$	Let $P = [\mathbf{x}_1 \mid \mid \mathbf{x}_n]$, then	For any $M \in \mathbb{R}^{m \times n}$, MM^T and M^TM are symmetric and	i.e. each X _i corresponds to i th component of data	Apply $GS q_1,, q_n \leftarrow GS(a_1,, a_n)$ to build ONB
$(u_j)^{\perp}$ (orthogonal compliment)	U⊥V ⇔ U¹ = V and vice-versa	p = m those tail-zeros don't exist)		$AP = [\lambda_1 \mathbf{x}_1 \dots \lambda_n \mathbf{x}_n] = [\mathbf{x}_1 \dots \mathbf{x}_n] \operatorname{diag}(\lambda_1, \dots, \lambda_n) = PD$	For any $M \in \mathbb{R}^{m \times n}$, MM' and M' M are symmetric and positive semi-definite	i.e. random vector $X = [X_1,, X_n]^T$ models the data	$(\mathbf{q}_1, \dots, \mathbf{q}_n) \in \mathbb{R}^m$ for $C(A)$ For exams: more efficient to compute as
Uniquely decompose next $U_j \not\ni \mathbf{a}_{j+1} = \mathbf{v}_{j+1} + \mathbf{u}_{j+1}$	$Y \subseteq X \implies X^{\perp} \subseteq Y^{\perp}$ and $X \cap X^{\perp} = \{0\}$ Any $x \in \mathbb{R}^n$ can be uniquely decomposed into	$\frac{\overline{\operatorname{diag}}_{m \times n}(\mathbf{a}) + \operatorname{diag}_{m \times n}(\mathbf{b}) = \operatorname{diag}_{m \times n}(\mathbf{a} + \mathbf{b})}{\operatorname{Consider diag}_{n \times n}(c_1, \dots, c_n), q = \min(n, k)}, \text{ then}$	$\det \left(\mathbf{A} \cdot \mathbf{U} \mathbf{W} \mathbf{V}^{T}\right) = \det \left(\mathbf{W}^{-1} \cdot \mathbf{V}^{T} \mathbf{A}^{-1} \mathbf{U}\right) \det(\mathbf{W}) \det(\mathbf{A})$	=> if <u>P</u> ⁻¹ exists then	Singular Value Decomposition (SVD) &	Co variance matrix of Viscou(A) = 1 ATA	For exams: more eπicient to compute as $\mathbf{u}_{j+1} = \mathbf{a}_{j+1} - Q_j \mathbf{c}_j$
$ \mathbf{v}_{j+1} = P_j(\mathbf{a}_{j+1}) \in U_j \Rightarrow \text{discard it!!}$	$\mathbf{x} = \mathbf{x}_i + \mathbf{x}_k$ where $\mathbf{x}_i \in U$ and $\mathbf{x}_k \in U^{\perp}$	$\operatorname{diag}_{m \times n}(a_1, \dots, a_p) \operatorname{diag}_{n \times k}(c_1, \dots, c_q)$	Tricks for computing determinant	A=PDP ⁻¹ i.e. Ajis diagonalisable P=I _{EB} is change-in-basis matrix for basis	Singular Values	Co-variance matrix of \underline{X} is $Cov(A) = \frac{1}{m-1} A^T A$	1. Gather $Q_j = [\mathbf{q}_1 \dots \mathbf{q}_j] \in \mathbb{R}^{m \times j}$ all-at-once
$ \mathbf{u}_{j+1} - \mathbf{v}_{j}(\mathbf{u}_{j+1}) = V_{\perp j}(\mathbf{u}_{j+1}) \in (U_j)^{\perp}$ => we're after this!!	For matrix $A \in \mathbb{R}^{m \times n}$ and for row-space $R(A)$,	$= \operatorname{diag}_{m \times k}(a_1 c_1, \dots, a_p) \operatorname{diag}_{n \times k}(c_1, \dots, c_q)$ $= \operatorname{diag}_{m \times k}(a_1 c_1, \dots, a_p c_p, 0, \dots, 0) = \operatorname{diag}_{n \times k}(c_1, \dots, c_q)$	If block-triangular matrix then apply	$B = \langle \mathbf{x}_1, \dots, \mathbf{x}_n \rangle$ of eigenvectors	Singular Value Decomposition of $A \in \mathbb{R}^{m \times n}$ is any decomposition of the form $A = USV^T$, where	$(A^T A)_{ij} = (A^T A)_{ji} = Cov(X_i, X_j)$	2. Compute $\mathbf{c}_j = [\mathbf{q}_1 \cdot \mathbf{a}_{j+1}, \dots, \mathbf{q}_j \cdot \mathbf{a}_{j+1}]^T \in \mathbb{R}^j$
	column-space $C(A)$ and null space $\ker(A)$ $R(A)^{\perp} = \ker(A)$ and $C(A)^{\perp} = \ker(A^{T})$	Where $r = \min(p, q) = \min(m, n, k)$ and	$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A) \det(B)$	If A = F _{EE} is transformation-matrix of linear map f then F _{EE} = I _{EB} F _{BB} I _{BE}	decomposition of the form $A = USV^T$ where $Orthogonal\ U = [u_1 u_m] \in \mathbb{R}^{m \times m}$ and	v ₁ ,, v _r (columns of V) are principal axes of A	all-at-once
Let $\underline{q_{j+1}} = \hat{\mathbf{u}}_{j+1} \Longrightarrow$ we have next ONB $(\underline{q_1,, q_{j+1}})$ for $U_{j+1} \Longrightarrow$ start next iteration	$R(A)^{\perp} = \ker(A)$ and $C(A)^{\perp} = \ker(A')$ Any $\mathbf{b} \in \mathbb{R}^{m}$ can be uniquely decomposed into	$s \in \mathbb{R}^{S}$, $s = \min(m, k)$	If close to triangular matrix apply EROs/ECOs to get it		$V = [\mathbf{v}_1 \mid \mid \mathbf{v}_n] \in \mathbb{R}^{n \times n}$	Let $\mathbf{w} \in \mathbb{R}^n$ be some unit-vector \Rightarrow let $\alpha_i = \mathbf{r}_i \cdot \mathbf{w}$ be the	3. Compute $Q_j c_j \in \mathbb{R}^m$ and subtract from a_{j+1}
$\mathbf{u}_{j+1} = (\mathbf{I}_m - Q_j Q_j^T) \mathbf{a}_{j+1} = \mathbf{a}_{j+1} - Q_j \mathbf{c}_j$ where	1 1 1 1 1 1 0(a) 1 1 (aT)	Inverse of square-diagonals => $diag(a_1,, a_n)^{-1} = diag(a_1^{-1},, a_n^{-1})$, i.e. diagonals	there, then its just product of diagonals If Cholesky/LU/QR is possible and cheap then do it,	Spectral theorem: if \underline{A} is Hermitian then \underline{P}^{-1} exists: $\ \mathbf{f} \mathbf{x}_i, \mathbf{x}_j\ _{\mathbf{A}}$ associated to different eigenvalues then	$S = \operatorname{diag}_{m \times n}(\sigma_1, \dots, \sigma_p)$ where $\underline{p} = \min(m, n)$ and $\sigma_1 \ge \cdots \ge \sigma_p \ge 0$	projection/coordinate of sample rj onto w	all-at-once
"J+1 - (*m - 'g) 'g J*J+1 - "j+1 - "G 'G Where	$b = b_i + b_k$, where $b_i \in R(A)$ and $b_k \in ker(A)$	diag $(a_1,, a_n)^{-1}$ = diag $(a_1^{-1},, a_n^{-1})$ i.e. diagonals cannot be zero (division by zero undefined)	then apply $ AB = A B $	$ \mathbf{x}_i \perp \mathbf{x}_j '$	· · · · · · · · · · ·		Can now rewrite $\underline{\mathbf{a}_j} = \sum_{i=1}^{J} (\mathbf{q}_i \cdot \mathbf{a}_j) \mathbf{q}_i = \mathbf{Q}_j \mathbf{c}_j$
$c_i = [q_1 \cdot a_{i+1}, \dots, q_i \cdot a_{i+1}]^T$							I————

Choose $Q = Q_n = [\mathbf{q}_1 \mid \mid \mathbf{q}_n] \in \mathbb{R}^{m \times n}$ notice its	*proj _{Lu} = uu ^T and proj _{Pu} = I _n -uu ^T =>	$\begin{vmatrix} \frac{\partial^n R^{*\cdots*n_1} f}{\partial x_i^n R \dots \partial x_i^{n_1}} &= \partial^n_{i_R} \dots \partial^n_{i_1} f &= f^{(n_1, \dots, n_R)}_{i_1 \dots i_R} \end{vmatrix}$	•, o, o, o, inner-product, back-substitution w/ triangular systems, are backwards stable	$fl(\lambda A) = \lambda A + E; E _{ij} \le \lambda A _{ij} \in mach$	Stability depends on growth-factor $\rho = \frac{\max_{i,j} u_{i,j} }{ u_{i,j} }$	Rayleigh quotient for <u>Hermitian</u> $A = A^{\dagger}$ is	Similar to to Gram-Schmidt (but different inner-product)
semi-orthogonal since Q ^T Q = I _n	H _u = proj _{Pu} - proj _{Lu}	i _k i ₁	If backwards stable \tilde{f} and f has condition number	$\begin{aligned} & fl(\mathbf{A} + \mathbf{B}) = (\mathbf{A} + \mathbf{B}) + \mathcal{E}; \mathcal{E} _{ij} \leq \mathbf{A} + \mathbf{B} _{ij} \in_{\text{mach}} \\ & + fl(\mathbf{A}\mathbf{B}) = \mathbf{A}\mathbf{B} + \mathcal{E}; \mathcal{E} _{ij} \leq n \in_{\text{mach}} (\mathbf{A} \mathbf{B})_{ij} + O(\in_{\text{mach}}^{2}) \end{aligned}$	max _{i,j} a _{i,j}	$R_A(\mathbf{x}) = \frac{\mathbf{x}^{\dagger} A \mathbf{x}}{\mathbf{x}^{\dagger} \mathbf{x}}$	$ (p^{(0)},, p^{(n-1)}) $ and $(r^{(0)},, r^{(n-1)}) $ are bases for
*Notice \Rightarrow $\mathbf{a}_j = Q_j \mathbf{c}_j = \mathbf{Q}[\mathbf{q}_1 \cdot \mathbf{a}_j,, \mathbf{q}_j \cdot \mathbf{a}_j, 0,, 0]^T = \mathbf{Q}\mathbf{r}_j$ *Let $R = [\mathbf{r}_1 \mathbf{r}_n] \in \mathbb{R}^{n \times n} \Rightarrow$	*Visualize as preserving component in Pu then flipping component in Lu	Its an \underline{N} -th order partial derivative where $\underline{N} = \sum_{k} n_{k}$ $\nabla f = [\partial_{1} f,, \partial_{n} f]^{T}$ is gradient of \underline{f} => $(\nabla f)_{i} = \frac{\partial f}{\partial \mathbf{x}_{i}}$	$ \underline{\kappa(x)} $ then relative error $ \underline{\ \tilde{f}(x)-f(x)\ } = O(\kappa(x)\varepsilon_{mach})$	(AB)=AB+E, [E][j] Shemach ([A[[B]])[j]+O(emach)	⇒ for partial pivoting $ρ ≤ 2^{m-1}$ $ U = O(ρ A) ⇒ \tilde{L}\tilde{U} = \tilde{P}A + δA , \frac{ δA }{ A } = O(ρε_{machine})$	Eigenvectors are stationary points of RA	QR Algorithm to find Schur decomposi-
$\begin{bmatrix} \mathbf{q}_1^T \mathbf{a}_1 & \dots & \mathbf{q}_1^T \mathbf{a}_n \end{bmatrix}$	-Н _и is involutory, orthogonal and symmetric, i.e.		Accuracy, stability, backwards stability are	Taylor series about $\underline{a} \in \mathbb{R}$ is	$\frac{\ U\ = O(\rho \ A\)}{\ A\ } \Rightarrow \underline{LU = PA + \delta A}, \frac{\ A\ }{\ A\ } = O(\rho \epsilon_{\text{machine}})$ $\Rightarrow \text{ only backwards stable if } \rho = O(1)$	$R_A(x)$ is closest to being like eigenvalue of x , i.e. $R_A(x) = \operatorname{argmin} Ax - \alpha x _2$	tion A = QUQ [†]
$A = QR = Q$ $\begin{bmatrix} 0 & q_n^T a_n \end{bmatrix}$, notice its	$H_{\boldsymbol{u}} = H_{\boldsymbol{u}}^{-1} = H_{\boldsymbol{u}}^{T}$	$ \underline{\nabla^T f} = (\nabla f)^T $ is $\underline{\text{transpose}}$ of $\underline{\nabla f}$, i.e. $\underline{\nabla^T f}$ is $\underline{\text{row vector}}$	norm-independent for fin-dim X, Y	$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (x-a)^k + O((x-a)^{n+1})$ as $x \to a$	- Only backwards stable ii p=0(1)	$R_A(x) - R_A(v) = O(x - v ^2)$ as $x \to v$ j where v j is	Any $\underline{A \in \mathbb{C}^{m \times m}}$ has Schur decomposition $\underline{A} = QUQ^{\dagger}$
upper-triangular	Modified Gram-Schmidt	$D_{\mathbf{u}} f(\mathbf{x}) = \lim_{\delta \to 0} \frac{f(\mathbf{x} + \delta \mathbf{u}) - f(\mathbf{x})}{\delta}$ is	Big-O meaning for numerical analysis in complexity analysis $f(n) = O(g(n)) as n \to \infty$	Need $\underline{a=0} = f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^{k} + O(x^{n+1})$ as	Full pivoting is PAQ = LU finds largest entry in bottom-right submatrix	eigenvector	Q is unitary, i.e. $Q^{\dagger} = Q^{-1}$ and upper-triangular U
Full OR Decomposition	Go check <u>Classical GM</u> first, as this is just an alternative computation method	directional-derivative of \underline{f}	But in numerical analysis $f(\varepsilon) = O(g(\varepsilon))$ as $\varepsilon \to 0$, i.e. $\limsup_{\varepsilon \to 0} f(\varepsilon) / g(\varepsilon) < \infty$	x → 0	Makes it pivot with row/column swaps before	Power iteration: define sequence $\frac{b^{(k+1)}}{ \mathbf{a}\mathbf{b}(k) } = \frac{A\mathbf{b}^{(k)}}{ \mathbf{a}\mathbf{b}(k) }$	Diagonal of <u>U</u> J contains eigenvalues of <u>A</u> J
+Consider full-rank $A = [a_1 a_n] \in \mathbb{R}^{m \times n} (\underline{m \ge n}),$	Let $P_{\perp} \mathbf{q}_j = \mathbf{I}_m - \mathbf{q}_j \mathbf{q}_j^T$ be projector onto hyperplane	It is rate-of-change in direction <u>u</u> , where <u>u</u> ∈ R ⁿ is unit-vector	i.e. ∃C,δ>0 s.t. <u>∀c.</u>], we have	te.g. $(1+\epsilon)^p = \sum_{k=0}^n {p \choose k} \epsilon^k + O(\epsilon^{n+1})$	very expensive O(m ³) search-ops, partial pivoting	Power iteration: define sequence D. Ab(k)	Algorithm 1 Basic QR iteration
i.e. a ₁ ,,a _n ∈ R ^m are linearly independent Apply QR decomposition to obtain:	$(Rq_j)^{\perp}$, i.e. orthogonal compliment of line Rq_j	$\frac{D_{\mathbf{u}}f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \ \nabla f(\mathbf{x})\ \ \mathbf{u}\ \cos(\theta)\ \Rightarrow D_{\mathbf{u}}f(\mathbf{x})\ \\ \text{maximized when } \cos\theta = 1$	$\begin{array}{c} 0 < \ \varepsilon\ < \delta \implies \ f(\varepsilon)\ \le C \ g(\varepsilon)\ \\ O(g) \text{ is set of functions} \end{array}$	$e.g.(1+\epsilon)^{p} = \sum_{k=0}^{n} \frac{p!}{k!(p-k)!} \epsilon^{k} + O(\epsilon^{n+1}) $ as $\epsilon \to 0$	only needs $O(m^2)$	with initial $b^{(0)}$ s.t. $\ b^{(0)}\ = 1$	1: for $k = 1, 2, 3,$ do
ONB $(\mathbf{q}_1, \dots, \mathbf{q}_n) \in \mathbb{R}^m$ for $C(A)$	-Notice: $P_{\perp j} = I_m - Q_j Q_j^T = \prod_{i=1}^{j} (I_m - q_i q_i^T) = \prod_{i=1}^{j} P_{\perp} q_i$	i.e. when x , u are parallel \Rightarrow hence $\nabla f(x)$ is direction	$\frac{ g(g) }{\{f: \limsup_{\epsilon \to 0} f(\epsilon) / g(\epsilon) < \infty\}}$	Elementary Matrices	Metric spaces & limits	Assume dominant $\lambda_1; x_1$ exist for \underline{A} and that $\text{proj}_{x_1} (b^{(0)}) * 0$	2: $A^{(k-1)} = Q^{(k-1)}R^{(k-1)}$ 3: $A^{(k)} = R^{(k-1)}Q^{(k-1)}$
Seni-ordiogonal Q1 = [q1] qn] e.k.		of max. rate-of-change		Identity $I_n = [e_1 \dots e_n] = [e_1; \dots; e_n]$ has	Metrics obey these axioms $d(x, x) = 0 \mid x \neq y \implies d(x, y) > 0 \mid d(x, y) = d(y, x) \mid$	Under above assumptions.	4: end for
*Compute basis extension to obtain remaining	-Re-state: $\mathbf{u}_{j+1} = (\mathbf{I}_m - Q_j Q_j^T) \mathbf{a}_{j+1} = \mathbf{v}_{j+1}$	f has local minimum at x_{loc} if there's radius $r > 0$ js.t.	Smallness partial order $O(g_1) \leq O(g_2)$ defined by set-inclusion $O(g_1) \subseteq O(g_2)$	elementary vectors e ₁ ,,e _n for rows/columns Row/column switching: permutation matrix P _{ii}	$d(x,z) \le d(x,y) + d(y,z)$	$\mu_{k} = R_{A} \left(\mathbf{b}^{(k)} \right) = \frac{\mathbf{b}^{(k)} + \mathbf{A} \mathbf{b}^{(k)}}{\mathbf{b}^{(k)} + \mathbf{b}^{(k)}}$ converges to dominant	For $\underline{A} \in \mathbb{R}^{m \times m}$ leach iteration $\underline{A}^{(k)} = \underline{Q}^{(k)} \underline{R}^{(k)}$ produces
$\mathbf{q}_{n+1}, \dots, \mathbf{q}_m \in \mathbb{R}^m$ where $\langle \mathbf{q}_1, \dots, \mathbf{q}_m \rangle$ is ONB for \mathbb{R}^m	$\mathbf{u}_{j+1} = \left(\prod_{i=1}^{j} P_{\perp \mathbf{q}_{i}}\right) \mathbf{a}_{j+1} = \left(P_{\perp \mathbf{q}_{j}} \cdots P_{\perp \mathbf{q}_{1}}\right) \mathbf{a}_{j+1}$	$\forall \mathbf{x} \in B[r; \mathbf{x}_{loc}]$ we have $f(\mathbf{x}_{loc}) \leq f(\mathbf{x})$	i.e. as $\underline{\epsilon} \to 0$, $g_1(\underline{\epsilon})$ goes to zero faster than $g_2(\underline{\epsilon})$ Roughly same hierarchy as complexity analysis but	obtained by switching ei and ej in In (same for	For metric spaces, mix-and-match these	B(w) B(w)	orthogonal $Q^{(k)^T} = Q^{(k)^{-1}}$
	Projectors P _{1 q1} ,,P _{1 qj} are iteratively applied to	f has global minimum x_{glob} if $\forall x \in \mathbb{R}^n$ we have $f(x_{glob}) \le f(x)$	flipped (some don't fit the pattern)	rows/columns) †Applying P _{ij} from left will swap rows, from right will	infinite/finite limit definitions: $ \frac{1}{\lim_{X\to+\infty} f(x)=+\infty} \iff \forall r \in \mathbb{R}, \exists N \in \mathbb{N}, \forall x > N: f(x) > r $	$\frac{h_1}{(b_k)}$ converges to some dominant x_1 jassociated with	$A^{(k+1)} = R^{(k)}Q^{(k)} = (Q^{(k)}^TQ^{(k)})R^{(k)}Q^{(k)}$ means
Let $Q_2 = [\mathbf{q}_{n+1} \dots \mathbf{q}_m] \in \mathbb{R}^{m \times (m-n)}$, let $Q = [Q_1 Q_2] \in \mathbb{R}^{m \times m}$, let $R = [R_1; 0_{m-n}] \in \mathbb{R}^{m \times n}$	a_{j+1} removing its components along q_1 then along q_2 and so on	A local minimum satisfies optimality conditions:	e.g. \dots , $O(\varepsilon^3) < O(\varepsilon^2) < O(\varepsilon) < O(1)$	swap columns		$\lambda_1 \Rightarrow Ab^{(k)} $ converges to $ \lambda_1 $	$= Q(k)^{T} A(k) Q(k)$
	Let $\mathbf{u}_{k}^{(j)} = \left(\prod_{i=1}^{j} P_{\perp} \mathbf{q}_{i}\right) \mathbf{a}_{k}$, i.e. \mathbf{a}_{k} without its	$\forall f(\mathbf{x}) = 0$ e.g. for $\underline{n} = 1$ its $f'(\mathbf{x}) = 0$	Maximum: $O(\max(g_1 , g_2)) = O(g_2) \iff O(g_1) \leq O(g_2)$	$P_{ij} = P_{ij}^T = P_{ij}^{-1}$, i.e. applying twice will undo it	$\lim_{X\to p} f(x) = L \iff \begin{cases} \forall \varepsilon > 0, \exists \delta > 0, \forall x \in X : \\ 0 < d_X(x,p) < \delta \implies d_Y(f(x),L) < \varepsilon \end{cases}$	If $\operatorname{proj}_{X_1}(b^{(0)}) = 0$ then $(b_k); (\mu_k)$ converge to second	$A^{(k+1)}$ is similar to $A^{(k)}$ Setting $A^{(0)} = A$ we get $A^{(k)} = (\tilde{Q}^{(k)})^T A \tilde{Q}^{(k)}$ where
$A = QR = [Q_1 Q_2] \begin{bmatrix} R_1 \\ 0_{m-n} \end{bmatrix} = Q_1 R_1$	components along $\mathbf{q}_1, \dots, \mathbf{q}_j$	$\nabla^2 f(\mathbf{x})$ is positive-definite, e.g. for $\underline{n} = 1$ its $\underline{f''(\mathbf{x})} > 0$	e.g. $O(\max(\epsilon^k, \epsilon)) = O(\epsilon)$	Row/column scaling : $D_i(\lambda)$ obtained by scaling e_i by λ in I_n (same for rows/columns)	Cauchy sequences, i.e. $\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \ge N : d(a_m, a_n) < \varepsilon$, converge in	dominant λ_2 ; x_2 instead If no dominant λ] (i.e. multiple eigenvalues of	$\tilde{Q}(k) = Q(0) \dots Q(k-1)$
Q is orthogonal , i.e. $Q^{-1} = Q^T$ so its a basis	Notice: $\mathbf{u}_j = \mathbf{u}_j^{(j-1)}$, thus $\mathbf{q}_j = \hat{\mathbf{u}}_j = \mathbf{u}_j^{(j-1)}/r_{jj}$ where	$\mathbf{H}(f) = \nabla^2 f = \mathbf{J}(\nabla f)^T$ is Hessian \Rightarrow $\mathbf{H}(f)_{ij} = \frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_i}$	Using functions $f_1,, f_n$ let $\Phi(f_1,, f_n)$ be formula	Applying P_{ij} from left will scale rows, from right will	You can manipulate matrix limits much like in real	maximum [λ] ∫ then ⟨b _R ⟩ will converge to linear combination of their corresponding	Under certain conditions QR algorithm converges to Schur decomposition
transformation		Interpret $F: \mathbb{R}^n \to \mathbb{R}^m$ as m functions $F_i: \mathbb{R}^n \to \mathbb{R}$	defining some function	scale columns	analysis, e.g. $\lim_{n\to\infty} (A^n B + C) = (\lim_{n\to\infty} A^n) B + C$	eigenvectors	
$-\operatorname{proj}_{C(A)} = Q_1 Q_1^T$, $\operatorname{proj}_{C(A)\perp} = Q_2 Q_2^T$ are	$r_{jj} = \left\ \mathbf{u}_{j}^{(j-1)} \right\ $ -Iterative step:	(one per output-component)	Then $\Phi(O(g_1),, O(g_n))$ is the <u>class of functions</u> $[\Phi(f_1,, f_n) : f_1 \in O(g_1),, f_n \in O(g_n)]$	$D_i(\lambda) = \text{diag}(1,, \lambda,, 1)$ so all diagonal properties apply, e.g. $D_i(\lambda)^{-1} = D_i(\lambda^{-1})$	Turn metric limit $\lim_{n\to\infty} x_n = L$ into real limit	Slow convergence if dominant λ ₁] not "very dominant"	We can apply shift $\mu^{(k)}$ at iteration k] $\Rightarrow A^{(k)} - \mu^{(k)} = O^{(k)} R^{(k)}$: $A^{(k+1)} = R^{(k)} O^{(k)} + \mu^{(k)} I$
orthogonal projections onto $C(A) \downarrow C(A)^{\perp} = \ker(A^{\top})$ respectively	$\begin{vmatrix} \mathbf{u}_k^{(j)} = \left(\mathbf{P}_{\perp \mathbf{q}_j} \right) \mathbf{u}_k^{(j-1)} = \mathbf{u}_k^{(j-1)} - \left(\mathbf{q}_j \cdot \mathbf{u}_k^{(j-1)} \right) \mathbf{q}_j \end{vmatrix}$	$\underline{\mathbf{J}(F)} = \left[\nabla^T F_1; \dots; \nabla^T F_m \right] \text{ is } \mathbf{Jacobian} \Rightarrow \underline{\mathbf{J}(F)_{ij}} = \frac{\partial F_i}{\partial \mathbf{x}_j}$	e.g. $\epsilon^{O(1)} = \{\epsilon^{f(\epsilon)} : f \in O(1)\}$	Row addition: $L_{ij}(\lambda) = L_{ij}(\lambda) = I_{ij}(\lambda)$ performs	$\lim_{n\to\infty} d(x_n, L) = 0$ to leverage real analysis	$\ \mathbf{b}^{(k)} - \alpha_k \mathbf{x}_1\ = O\left(\left\ \frac{\lambda_2}{\lambda_1}\right\ ^k\right) $ for phase factor	If shifts are good eigenvalue estimates then
Notice: $QQ^T = I_m = Q_1Q_1^T + Q_2Q_2^T$	$R \left(\frac{1}{2} \frac{1}{4} \right) / R R \left(\frac{3}{2} R \right) / \frac{3}{4}$ -i.e. each iteration j of MGS computes $P_{\perp} \mathbf{q}_{i}$ (and	Conditioning	General case:	$R_i \leftarrow R_i + \lambda R_j$ when applying from left	Bounded monotone sequences converge in R Sandwich theorem for limits in R > pick easy	$\alpha_k \in \{-1, 1\}$ it may alternate if $\lambda_1 < 0$	last column of $\tilde{Q}^{(k)}$ converges quickly to an
•Generalizable to <u>A ∈ C^{m×n}</u> by changing transpose to conjugate-transpose	projections under it) in one go	A problem is some $\underline{f}: X \to Y$ where \underline{X}, Y are normed vector-spaces	$ \frac{\Phi_1(O(f_1),, O(f_m)) = \Phi_2(O(g_1),, O(g_n))}{\Phi_1(O(f_1),, O(f_m)) \subseteq \Phi_2(O(g_1),, O(g_n))} \text{means}$	$\lambda e_i e_i^T$ is zeros except for λ in (i,j) th entry	upper/lower bounds	(λ ₁) ^k c ₁ . †, (α)	eigenvector Estimate μ ^(k) with Rayleigh quotient =>
Lines and hyperplanes in En(-Dn)	At start of iteration $j \in 1n$ we have ONB	A problem <i>instance</i> is f with fixed input x ∈ X, shortened to <i>just</i> "problem" (with x ∈ X, implied)	e.g. $e^{O(1)} = O(k^{\epsilon})$ means $\{e^{f(\epsilon)} : f \in O(1)\} \subseteq O(k^{\epsilon})$	$L_{ij}(\lambda)^{-1} = L_{ij}(-\lambda)$ both triangular matrices	$\frac{\lim_{n\to\infty} r^n = 0 \iff r < 1}{\lim_{n\to\infty} \sum_{i=0}^n ar^i = \frac{a}{1-r} \iff r < 1}$	10(1) (-1)	$\mu^{(k)} = (A_k)_{mm} = (\tilde{\mathbf{q}}_m^{(k)})^T A \tilde{\mathbf{q}}_m^{(k)} \text{ where } \tilde{\mathbf{q}}_m^{(k)} \text{ is } \underline{m} \text{ th}$
	$\mathbf{q}_1, \dots, \mathbf{q}_{j-1} \in \mathbb{R}^m$ and residual $\mathbf{u}_j^{(j-1)}, \dots, \mathbf{u}_n^{(j-1)} \in \mathbb{R}^m$	δx is small perturbation of $x_j \Rightarrow \delta f = f(x + \delta x) - f(x)$	not necessarily true Special case: $f = \Phi(O(g_1),, O(g_n))$ means	LU factorization w/ Gaussian elimina-	Iterative Techniques	b(k); x ₁ are normalized	column of $ ilde{Q}^{(k)}$
with standard basis $(e_1,, e_n) \in \mathbb{R}^n$ with standard origin $0 \in \mathbb{R}^n$	-Compute $r_{jj} = \left\ \frac{\mathbf{u}_{j}^{(j-1)}}{\mathbf{u}_{j}^{j}} \right\ \Rightarrow \mathbf{q}_{j} = \frac{\mathbf{u}_{j}^{(j-1)}}{r_{jj}} / r_{jj}$	A problem (instance) is: Well-conditioned if all small δx lead to small δf , i.e.	$f \in \Phi(O(g_1), \dots, O(g_n))$	Recall: you can represent EROs and ECOs as	Systems of Equations	$(A-\sigma I)$ has eigenvalues $\lambda - \sigma$	
	-For each $k \in (j+1)n$ compute $r_{jk} = \mathbf{q}_j \cdot \mathbf{u}_k^{(j-1)} = $	if K is small (e.g. 1] 10] 10 ²	e.g. $(\varepsilon+1)^2 = \varepsilon^2 + O(\varepsilon)$ means	transformation matrices R, C respectively	Let $A, R, G \in \mathbb{R}^{n \times n}$ where G^{-1} exists \Longrightarrow splitting	\Rightarrow power-iteration on $(A-\sigma I)$ has $\frac{\lambda_2-\sigma}{\lambda_1-\sigma}$	
A line $L = \mathbb{R} n + c$ is characterized by direction $n \in \mathbb{R}^n$ $(n \neq 0)$ and offset from origin $c \in L$	$\frac{\mathbf{u}_{k}^{(j)} = \mathbf{u}_{k}^{(j-1)} - r_{jk}\mathbf{q}_{j}}{\mathbf{q}_{j}}$	-Ill-conditioned if some small δx lead to large δf , i.e.	$ \underline{\epsilon} \mapsto (\epsilon + 1)^2 \in \{\epsilon^2 + f(\epsilon) : f \in O(\epsilon)\} $ not necessarily true	LU J factorization => finds A = LU J where L, U J are lower/upper triangular respectively	A=G+RJhelps iteration Ax=bJrewritten as x=Mx+cJwhere	Eigenvector guess => estimated eigenvalue	
It is a section of the second	-Next ONB $(\mathbf{q}_1,, \mathbf{q}_j)$ and next residual $\mathbf{u}_{i+1}^{(j)},, \mathbf{u}_n^{(j)}$	if <u>K</u> jis large (e.g. <u>10⁶</u> , <u>10¹⁶</u>)	Let $f_1 = O(g_1)$, $f_2 = O(g_2)$ and let $k \neq 0$ be a constant	Naive Gaussian Elimination performs	$ M = -G^{-1}R; c = -G^{-1}b $ Define $f(x) = Mx + c$ and sequence	Inverse (power-)iteration: perform power iteration on	
-c∈L is closest point to origin , i.e. c⊥n	-NOTE: for $j=1$ => $\mathbf{q}_1,, \mathbf{q}_{j-1} = \emptyset$ i.e. none yet	Absolute condition number $cond(x) = \hat{\kappa}(x) = \hat{\kappa}$ of \underline{f} at \underline{x} :	$ f_1 f_2 = O(g_1 g_2) f \cdot O(g) = O(fg) O(k \cdot g) = O(g) $ $ f_1 + f_2 = O(\max(g_1 , g_2)) $	$[I_m \mid A \mid I_n] \rightsquigarrow [R^{-1} \mid U \mid I_n]$ to get $AI_n = R^{-1} U$ using	$\mathbf{x}^{(k+1)} = f(\mathbf{x}^{(k)}) = M\mathbf{x}^{(k)} + \mathbf{c}$ with starting point $\mathbf{x}^{(0)}$	$(A-\sigma I)^{-1}$ to get $\lambda_{1,\sigma}$ closest to σ	
·If c≠λn]⇒ L]not vector-subspace of ℝ ⁿ] †i.e. 0∉L¦i.e. L]doesn't go through the origin	By end of iteration $j = n$, we have ONB	$\hat{\kappa} = \lim_{\delta \to 0} \sup_{\ \delta x\ \le \delta} \frac{\ \delta f\ }{\ \delta x\ }$	\Rightarrow if $g_1 = g = g_2$ then $f_1 * f_2 = O(g)$	only row addition R ⁻¹ , i.e. inverse EROs in reversed order, is	Limit of (x_R) is fixed point of $f \Rightarrow$ unique fixed point of f is solution to $Ax = b$	$(A-\sigma I)^{-1}$ has eigenvalues $(\lambda-\sigma)^{-1}$ so power iteration will yield largest $(\lambda_{1,\sigma}-\sigma)^{-1}$	
LJis affine-subspace of Rn	$(\mathbf{q}_1, \dots, \mathbf{q}_n) \in \mathbb{R}^m$	=> for $\underline{\text{most problems}}$ simplified to $\hat{\kappa} = \sup_{\delta X} \frac{\ \delta f\ }{\ \delta x\ }$	Floating-point numbers Consider base/radix β≥2 (typically 2) and precision	lower-triangular so L=R ⁻¹	If - is consistent norm and M < 1 then (x _k)	i.e. will yield smallest $\lambda_{1,\sigma} - \sigma$ i.e. will yield $\lambda_{1,\sigma}$	
•If c= λn, i.e. L= Rn => L is vector-subspace of R ⁿ +i.e. 0∈ L , i.e. L goes through the origin	$-A = [a_1 a_n] = [q_1 q_n]$ $r_{11} r_{1n}$ $r_{1n} = QR$	If $\underline{Jacobian} \ \underline{J}_f(x)$ exists then $\hat{\kappa} = \ \underline{J}_f(x)\ $ where $\underline{Matrix norm} \ \ -\ $ induced by $\underline{Matrix} \ norms \ $	t≥1](24]or 53]for IEEE single/double precisions)	Algorithm 1 Gaussian elimination	converges for any x(0) (because Cauchy-completeness)	closest to g	
Ljhas dim(L) = 1 and orthonormal basis (ONB) {n̂}	0 r _{nn} corresponds to thin QR decomposition	Relative condition number $\kappa(x) = \kappa of f at x_j is$	Floating-point numbers are discrete subset $F = \{ (-1)^S (m/\beta^t) \beta^e \mid 1 \le m \le \beta^t, s \in \mathbb{B}, m, e \in \mathbb{Z} \} $	2: for $k = 1$ to $m - 1$ do	-We want to find M < 1 and easy to compute M; c -Stopping criterion usually the relative residual	$\ \mathbf{b}^{(k)} - \alpha_k \mathbf{x}_{1,\sigma}\ = O\left(\left\ \frac{\lambda_{1,\sigma} - \sigma}{\lambda_{2,\sigma} - \sigma}\right\ ^{R}\right) \text{ where } \mathbf{x}_{1,\sigma}$	
A hyperplane $P = (Rn)^{\perp} + c = \{x + c \mid x \in R^n, x \perp n\}$ is	Where $A \in \mathbb{R}^{m \times n}$ is full-rank, $\mathbb{Q} \in \mathbb{R}^{m \times n}$ is	$\kappa = \lim_{\delta \to 0} \sup_{\ \delta x\ \le \delta} \left(\frac{\ \delta f\ }{\ f(x)\ } / \frac{\ \delta x\ }{\ x\ } \right)$	sjis sign-bit, m/β ^t is mantissa, ejis exponent (8]-bit	3: for $j = k + 1$ to m do 4: $\ell_{j,k} = u_{j,k}/u_{k,k}$	b-Ax ^(k)	corresponds to $\lambda_{1,\sigma}$ and $\lambda_{2,\sigma}$ is 2nd-closest to σ	
={x∈R" x·n=c·n}	semi-orthogonal, and R∈R ^{n×n} is upper-triangular	=> for most problems simplified to	for single, 11 bit for double) Equivalently, can restrict to $\beta^{t-1} \le m \le \beta^t - 1$ for unique	5: $u_{j,k:m} = u_{j,k:m} - \ell_{j,k}u_{k,k:m}$ 6: end for		-Efficiently compute eigenvectors for known eigenvalues σ	
origin c∈P	Classical vs. Modified Gram-Schmidt These algorithms both compute thin	$\kappa = \sup_{\delta x} \left(\frac{\ \delta f\ }{\ f(x)\ } / \frac{\ \delta x\ }{\ x\ } \right)$	mjand ej	7: end for	Assume Afs diagonal is non-zero (w.l.o.g.	Eigenvalue guess ⇒ estimated eigenvector	
It represents an (n-1) dimensional slice of the	thin QR decomposition	[c(x)	F∈R] is idealized (ignores over/underflow), so is countably infinite and self-similar (i.e. F = βF)	The pivot element is simply diagonal entry $u_{kk}^{(k-1)}$	permute/change basis if isn't) then $A = D + L + U$; where D is diagonal of A , L , U are strict lower/upper triangular	Algorithm 3 Inverse iteration 1: for $k = 1, 2, 3,$ do	
It is customary that: In is a unit vector, i.e. n = n = 1		If <u>Jacobian</u> $J_f(x)$ exists then $\kappa = \frac{\ x\ _f \ x\ _f}{\ f(x)\ _f \ x\ _f}$ More important than $\hat{\kappa}$ for numerical analysis	For all $x \in \mathbb{R}$ there exists $f(x) \in \mathbb{F}$ s.t.	fails if $u_{kk}^{(k-1)} \approx 0$	parts of Al	2: $\hat{x}^{(k)} = (A - \sigma I)^{-1}x^{(k-1)}$ 3: $x^{(k)} = \hat{x}^{(k)}/\max(\hat{x}^{(k)})$	
-c∈P]is closest point to origin, i.e. c=λn	1: for $j = 1$ to n do 3: end for	Matrix condition number Cond(A) = K(A) = A A^{-1}	$ x-fl(x) \le \epsilon_{mach} x $ Equivalently $fl(x) = x(1+\delta), \delta \le \epsilon_{mach}$	$+\underline{\tilde{L}\tilde{U}} = A + \delta A$, $\underline{\ \tilde{L}\ \cdot \ \tilde{U}\ } = O(\epsilon_{mach})$ only backwards	Jacobi Method:	4: $\lambda^{(k)} = (x^{(k)})^T A x^{(k)}$	
-With those $\Rightarrow P = \{x \in \mathbb{R}^n \mid x \cdot \mathbf{n} = \lambda\}$	2: $uj = aj$ 4: for $j = 1$ to n do 3: for $j = 1$ to $j - 1$ do 5: $r_{ij} = u_j _2$	=> comes up so often that has its own name A ∈ € ^{m×m} is well-conditioned if κ(A) is small,	Machine epsilon $\epsilon_{\text{machine}} = \epsilon_{\text{mach}} = \frac{1}{2} \beta^{1-t}$	stable if L · U ≈ A	$\frac{G = D; R = L + U}{ c } \Rightarrow \frac{M = -D^{-1}(L + U); c = D^{-1}b }{ c }$	5: end for	
·If <u>c·n≠0</u> => <u>P</u> not vector-subspace of <u>R</u> ⁿ +i.e. <u>0∉P</u> , i.e. <u>P</u> doesn't go through the origin	4: $r_{ij} = q_i^* a_j$ 6: $q_j = u_j / r_{jj}$ 5: $u_j = u_j - r_{ij} q_i$ 7: for $k = j + 1$ to n do 6: end for 8: $r_{ij} = q_i^* u_i$	ill-conditioned if large	is maximum relative gap between FPs Half the gap between 1 and next largest FP	Work required: $\sim \frac{2}{3} m^3$ flops $\sim O(m^3)$	$\frac{\mathbf{x}_{i}}{\mathbf{x}_{i}} = \frac{\mathbf{x}_{i}}{\mathbf{x}_{i}} \left(\mathbf{v}_{i} - \mathbf{z}_{j \neq i} \mathbf{x}_{ij} \mathbf{x}_{j} \right) - \mathbf{x}_{i}$ only needs	Can reduce matrix inversion $O(m^3)$ to $O(m^2)$ by pre-factorization	
$-P$ is affine-subspace of \mathbb{R}^n $+ \mathbb{I} \cdot \mathbb{C} \cdot \mathbb{n} = 0$, i.e. $P = (\mathbb{R}^n)^{\perp} = P$ is vector-subspace of	7: $r_{ij} = u_i _2$ 9: $u_k = u_k - r_{ik}q_i$	$\frac{\kappa(\mathbf{A}) = \kappa(\mathbf{A}^{-1})}{\kappa(\mathbf{A}) = \kappa(\mathbf{A})} \underbrace{\ \kappa(\mathbf{A}) = \kappa(\mathbf{A})\ }_{\mathbf{A}} \underbrace{\ \cdot\ _{\mathbf{A}} = \ \cdot\ _{\mathbf{A}}}_{\mathbf{A}} \implies \kappa(\mathbf{A}) = \frac{\sigma_{\mathbf{A}}}{\sigma_{\mathbf{B}}}$	Hair the gap between 1 and next largest FP $2^{-24} \approx 5.96 \times 10^{-8}$ and $2^{-53} \approx 10^{-16}$ for single/double	Solving $\underline{Ax = LUx}$ Jis $\frac{2}{3}m^3$ flops (back substitution is $O(m^2)$)	$ \mathbf{b}_i; \mathbf{x}^{(k)}; A_{i\star} \Rightarrow \text{row-wise parallelization}$	Nonlinear Systems of Equations	
R ⁿ	8: $q_j = u_j/r_{jj}$ 10: end for 9: end for 11: end for	For $\underline{\mathbf{A}} \in \mathbb{C}^{m \times n}$ the problem $f_{\mathbf{A}}(x) = \mathbf{A}x$ has		NOTE: Householder triangularisation requires $\sim \frac{4}{3} m^3$	Gauss-Seidel (G-S) Method:	Recall that $\nabla f(\mathbf{x})$ is direction of max. rate-of-change $\nabla f(\mathbf{x})$	
	Computes at j th step: -Classical GS => j th column of Q and the j th column	$\kappa = \ \mathbf{A}\ \frac{\ \mathbf{x}\ }{\ \mathbf{A}\mathbf{x}\ } \implies \text{if } \underline{\mathbf{A}}^{-1} \text{ exists then } \underline{\kappa \leq \text{Cond}(\mathbf{A})}$	FP arithmetic: let ∗,⊕ be real and floating counterparts of arithmetic operation		$G = D + L; R = U => M = -(D + L)^{-1} U; c = (D + L)^{-1} b$	<u>Idea:</u> Search for stationary point by gradient descent : $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$ for step length $\underline{\alpha}$	
		$\ \mathbf{A}\mathbf{x}\ $ If $\ \mathbf{A}\mathbf{x} = \mathbf{b}\ $ problem of finding \mathbf{x} Igiven \mathbf{b} is just	For $x, y \in F$ we have	Partial pivoting computes $PA = LU$ where P is a permutation matrix $\Rightarrow PP^T = I$ i.e. its orthogonal	$\frac{\mathbf{x}_{i}^{(k+1)} = \frac{1}{A_{ii}} \left(\mathbf{b}_{i} - \sum_{j=1}^{i-1} A_{ij} \mathbf{x}_{j}^{(k+1)} - \sum_{j=i+1}^{n} A_{ij} \mathbf{x}_{j}^{(k)} \right)}{\left((1 - 1)^{i} \right)}$		
Notice <u>L</u> = Rn Jand P = (Rn) in are orthogonal compliments, so:	-Modified GS ⇒ j th column of Q Jand the j th row of RJ	If $\underline{\mathbf{A}} = b$. Problem of finding x given \underline{b} is just $f_{\underline{\mathbf{A}}-1}(b) = \underline{\mathbf{A}}^{-1}b$ $\Rightarrow \kappa = \ \underline{\mathbf{A}}^{-1}\ \frac{\ \underline{b}\ }{\ \underline{x}\ } \le Cond(\underline{\mathbf{A}})$	$x \circledast y = fl(x * y) = (x * y)(1 * \epsilon), \delta \le \epsilon_{mach}$ Holds for any arithmetic operation $\circledast = \bullet, \bullet, \bullet, \bullet$	For each column il finds largest entry and row-swaps	Computing $\mathbf{x}_{i}^{(k+1)}$ needs \mathbf{b}_{i} ; $\mathbf{x}^{(k)}$; $\mathbf{A}_{i\star}$ and $\mathbf{x}_{j}^{(k+1)}$ for	If <u>AJ</u> is <u>positive</u> -definite, solving <u>Ax = b</u> and $\min_{\mathbf{X}} f(\mathbf{X}) = \frac{1}{2} \mathbf{X}^T A \mathbf{X} - \mathbf{X}^T \mathbf{b}$ are equivalent	
0 1 2	Both have flop (floating-point operation) count of $O(2mn^2)$	For $\underline{\mathbf{b} \in \mathbb{C}^m}$, the problem $f_{\underline{\mathbf{b}}}(A) = A^{-1}\underline{\mathbf{b}}$ (i.e. finding $\underline{\mathbf{x}}$ _1 in	Complex floats implemented pairs of real floats, so	to make it new pivot => Pj	j < i ⇒ lower storage requirements	$ \frac{\min_{\mathbf{X}} f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^t A \mathbf{x} - \mathbf{x}^t \mathbf{b}}{\text{Get iterative methods } \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)} \mathbf{p}^{(k)}} \text{ for } $	
proj _P = id _R n -proj _L = I _n - n̂n̂ ^T is orthogonal projection onto PJ(along LJ)	NOTE: Householder method has $2(mn^2 - n^3/3)$ flop	$\underline{Ax = b}$ has $\underline{\kappa} = A A^{-1} = \text{Cond}(A)$	above applies to complex ops as-well Caveat: $\epsilon_{\text{mach}} = \frac{1}{2} \beta^{1-t}$ must be scaled by factors		Successive over-relaxation (SOR):	step length $\underline{\alpha}^{(k)}$ and directions $\underline{p}^{(k)}$	
·L=im(proj _L)=ker(proj _P) and	count, but better numerical properties	Stability Given a problem $f: X \to Y$ an algorithm for f is	on the order of $2^{3/2}$, $2^{5/2}$ for \emptyset , \emptyset respectively	Result is $L_{m-1}P_{m-1}L_2P_2L_1P_1A=U$, where	$\frac{G = \omega^{-1} D + L; R = (1 - \omega^{-1}) D + U}{M = -(\omega^{-1} D + L)^{-1} ((1 - \omega^{-1}) D + U); c = -(\omega^{-1} D + L)^{-1} b}$	Conjugate gradient (CG) method: if $A \in \mathbb{R}^{n \times n}$	
P=ker(proj _L)=im(proj _P)	Recall: Q [†] Q=I _n => check for loss of orthogonality	$\tilde{f}: X \to Y$	(v. a. av.)	$\frac{L_{m-1}P_{m-1} \dots L_2P_2L_1P_1 = L'_{m-1} \dots L'_1P_{m-1} \dots P_1}{\text{Setting } L = (L'_{m-1} \dots L'_1)^{-1} \mid P = P_{m-1} \dots P_1 \mid \text{gives}}$	$M = -(\omega^{-1}D + L)^{-1}((1 - \omega^{-1})D + U); \mathbf{c} = -(\omega^{-1}D + L)^{-1}\mathbf{b}$ $\downarrow_{\mathbf{v}}(k+1) = \frac{\omega}{A_{ij}} \left(\mathbf{b}_{i} - \sum_{j=1}^{i-1} A_{ij} \mathbf{x}_{j}^{(k+1)} - \sum_{j=i+1}^{n} A_{ij} \mathbf{x}_{j}^{(k)} \right) \Big _{\text{for}}$	symmetric then $\langle \mathbf{u}, \mathbf{v} \rangle_{A} = \mathbf{u}^{T} A \mathbf{v}$ is an inner-product	
	with $\ \mathbf{I}_n - Q^{\dagger} Q\ = \text{loss}$ Classical GS => $\ \mathbf{I}_n - Q^{\dagger} Q\ \approx \text{Cond}(A)^2 \in_{\text{mach}}$	Input $\underline{x \in X}$ is first rounded to $\underline{fl(x)}$ i.e. $\underline{\tilde{f}(x)} = \underline{\tilde{f}(fl(x))}$ Absolute error $\Rightarrow \ \tilde{f}(x) - f(x)\ $	$\approx (x_1 + \dots + x_n) + \sum_{i=1}^n x_i \left(\sum_{j=i}^n \delta_j \right)^{i, \delta_j \le \epsilon_{\text{mach}}}$	PA=LU	$ + \mathbf{x}_{i}^{(k+1)} = \overline{A_{ii}} \left(\begin{array}{ccc} \mathbf{v}_{i} - Z_{j=1} & A_{ij} & \mathbf{x}_{j} \\ + (1 - \omega) \mathbf{x}_{i}^{(k)} \end{array} \right) $ for	GC chooses p(k) that are conjugate w.r.t. Al i.e.	
Householder Maps: reflections	-Modified GS $\Rightarrow \ \mathbf{I}_{\mathbf{I}_{\mathbf{I}}} - \mathbf{Q}^{\dagger} \mathbf{Q} \ \approx \text{Cond}(\mathbf{A}) \in \text{mach} \ $	relative error $\Rightarrow \frac{\ \tilde{f}(x) - f(x) \ }{\ f(x) \ }$	$(x_1 \otimes \cdots \otimes x_n) \approx (x_1 \times \cdots \times x_n)(1 + \varepsilon), \varepsilon \le 1.06(n-1)\varepsilon_{\text{mach}}$ $f((\sum x_i y_i) = \sum x_i y_i(1 + \varepsilon_i))$ where	Algorithm 2 Gaussian elimination with partial pivoting 1: $U = A, L = I, P = I$	relaxation factor $\omega > 1$	$(\mathbf{p}^{(i)}, \mathbf{p}^{(j)})_A = 0$ for $i \neq j$	
Two points x , y ∈ E ⁿ are reflections w.r.t hyperplane	NOTE: Householder method has $\ \mathbf{I}_n - \mathbf{Q}^{\dagger}\mathbf{Q}\ \approx \epsilon_{\text{mach}}$		$1+\epsilon_i = (1+\delta_i) \times (1+\eta_i) \cdots (1+\eta_n)$ and $ \delta_j , \eta_i \le \epsilon_{mach}$	2: for $k = 1$ to $m - 1$ do 3: $i = \operatorname{argmax} u_{i,k} $		And chooses $\underline{\alpha}^{(k)}$ s.t. residuals $\underline{r}^{(k)} = -\nabla f(\underline{x}^{(k)}) = b - A\underline{x}^{(k)}$ are orthogonal	
P=(Rn) [±] +c if: 1. The translation xv=v-x lis parallel to normal n. i.e.	Multivariate Calculus	\underline{f} is accurate if $\underline{\forall x \in X}$. $\underline{\frac{\ \widehat{f}(x) - f(x)\ }{\ f(x)\ }} = O\left(\varepsilon_{\text{mach}}\right)$	$1+\epsilon_i \approx 1+\delta_i + (\eta_i + \dots + \eta_n)$	4: $u_{k,k:m} \stackrel{i \ge k}{\leftrightarrow} u_{i,k:m}$	If A Jis strictly row diagonally dominant then Jacobi/Gauss-Seidel methods converge; A Jis strictly	$ k=0 \Rightarrow \mathbf{p}^{(0)} = -\nabla f(\mathbf{x}^{(0)}) = \mathbf{r}^{(0)}$	
xy=λn	Consider $f: \mathbb{R}^n \to \mathbb{R}$: When clear write \underline{i} -th component of input as \underline{i} instead	\tilde{f} is stable if $\forall x \in X$, $\exists \tilde{x} \in X$ s.t. $\ \tilde{f}(x) - f(\tilde{x})\ $ and $\ \tilde{f}(x) - f(\tilde{x}$	$ fl(x^Ty)-x^Ty \le \sum x_iy_i \epsilon_i $ Assuming $n\epsilon_{\text{mach}} \le 0.1 =>$	5: $\ell_{k,1:k-1} \leftrightarrow \ell_{i,1:k-1}$ 6: $p_{k,:} \leftrightarrow p_{i,:}$	row diagonally dominant if $ A_{ij} > \sum_{j \neq i} A_{ij} $	$ \underline{k \ge 1} \Rightarrow \mathbf{p}^{(k)} = \mathbf{r}^{(k)} - \sum_{i < k} \frac{\langle \mathbf{p}^{(i)}, \mathbf{r}^{(k)} \rangle_{\mathbf{A}}}{\langle \mathbf{p}^{(i)}, \mathbf{p}^{(i)} \rangle_{\mathbf{A}}} \mathbf{p}^{(i)}$	
2 Midpoint $m = 1/2(\mathbf{x} \cdot \mathbf{y}) \in P$ lies on P i.e. $\underline{m \cdot \mathbf{n}} = \mathbf{c} \cdot \mathbf{n}$ •Suppose $P_{\mathbf{u}} = (\mathbf{R}\mathbf{u})^{\perp}$ goes through the origin with unit	of xi	$\frac{\ \tilde{f}(x)-f(\tilde{x})\ }{\ f(\tilde{x})\ } = O\left(\varepsilon_{\text{mach}}\right) \text{ and } \frac{\ \tilde{x}-x\ }{\ x\ } = O\left(\varepsilon_{\text{mach}}\right)$ †i.e. nearly the right answer to nearly the right question	$ fl(x^Ty)-x^Ty \le \phi(n)\epsilon_{mach} x ^T y $, where $ x _i = x_i $	7: for $j = k + 1$ to m do 8: $\ell_{j,k} = u_{j,k}/u_{k,k}$	If A j is positive-definite then G-S and SOR $(\omega \in (0, 2))$ converge	(b) (b) I	
	of x_i -Level curve w.r.t. to $c \in R$ Jis all points s.t. $f(x) = c$ -Projecting level curves onto R^n gives f f	outer-product is stable	is vector and φ(n) is small function of n Summing a series is more stable if terms	9: $u_{j,k;m} = u_{j,k;m} - \ell_{j,k}u_{k,k;m}$ 10: end for	Eigenvalue Problems	$= \operatorname{argmin}_{\alpha} f(\mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}) = \frac{\mathbf{p}^{(k)} \cdot \mathbf{r}^{(k)}}{(\mathbf{p}^{(k)}, \mathbf{p}^{(k)})_A}$	
Householder matrix $H_{u} = I_{n} - 2uu^{T}$ is reflection w.r.t. hyperplane P_{u}	contour-map	\tilde{f} is backwards stable if $\forall x \in X$, $\exists \tilde{x} \in X$ s.t. $\tilde{f}(x) = f(\tilde{x})$	added in order of increasing magnitude	11: end for	If A Jis diagonalizable then eigen-decomposition is	Without rounding errors, CG converges in ≤nj	
Recall: let Lu = Ru	n_k th order partial derivative w.r.t i_k of, of n_1 th	and $\frac{\ \bar{x}-x\ }{\ x\ } = O(\epsilon_{\text{mach}})$	For FP matrices , let $ M _{ij} = M_{ij} $ i.e. matrix $ M $ of	Work required: $\sim \frac{2}{3} m^3$ flops $\sim O(m^3)$; results in $L_{ij} \le 1$	Dominant λ_1 ; x_1 are such that $ \lambda_1 $ is strictly largest	iterations	
	order partial derivative w.r.t i of f is:	i.e. exactly the right answer to nearly the right question, a subset of stability	absolute values of M	so <u> L = O(1)</u>	for which Ax = \lambda x		
		*					I and the second