Pierson Guthrey
pguthrey@iastate.edu

Contents

Pierson Guthrey
pguthrey@iastate.edu

Part I
Linear Algebra

Linear Algebra concerns vector spaces and linear maps between those vector spaces.
•Geometric interpretation: Vectors are a class of arrows defined by length and direction.
•Analytic Viewpoint: Vectors take the form

$$\vec{\mathbf{v}} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad v_i \in \mathbb{F}$$

with vector addition and scalar multiplication defined as such:

$$\alpha\vec{\mathbf{v}} + \vec{\mathbf{w}} = \begin{pmatrix} \alpha v_1 + w_1 \\ \alpha v_2 + w_2 \\ \vdots \\ \alpha v_n + w_n \end{pmatrix} \quad v_i, w_i \in \mathbb{F}, \alpha \in \mathbb{R}$$

1 Vector Spaces

A vector space V over a field $\mathbb{F}$ has two operations:

- Vector addition: $(u, v) \in V \to w \in V$
  Properties: Abelian (commutative), associative, invertible ($-v$), unit element (zero vector)
- Scalar Multiplication: $a \in \mathbb{R}, v \in V \to w \in V$
  Properties: Unit element (1), Distributive: Multiplication $\alpha(\beta v) = (\alpha\beta)v$, Distributive: Addition

Vector spaces are closed under both operations.
If we consider functions to be vectors, then the following are vector spaces

- $\mathbb{R}^n$ with scalars in $\mathbb{R}$, and $\mathbb{C}^n$ with scalars in $\mathbb{C}$
- $C(\Omega)$, $C^m(\Omega)$, $C^\infty(\Omega)$ are vector spaces
- $P^n \subset C^\infty(\Omega)$ is a vector space of polynomials of degree less than n
- $P \subset C^\infty(\Omega)$ is a vector space of polynomials of any degree
- $L^p(\Omega)$ with $1 \leq p \leq \infty$ are vector spaces Note that $d(f, g) = 0 \iff f = g$ a.e. since they can differ on a set of measure zero and still have the same integral, thus two such functions must be regarded as identical, so $f \sim g$ (equivalence relation) if $f = g$ a.e.
- The general solutions of PDEs are a vector space

1.1 Linear Combinations of vectors

$$\sum_{i=1}^{n} \alpha_i \vec{\mathbf{v}}_i$$

Vectors are linearly dependent if $\exists\, \alpha_i$ such that $\sum_{i=1}^{n} \alpha_i \vec{\mathbf{v}}_i = 0$ but not all $\alpha_i = 0$. (this includes any space that contains the zero vector)
Linearly independence: A set of vectors is linearly independent if

$$\sum_{i=1}^{n} \alpha_i \vec{\mathbf{v}}_i = 0 \Rightarrow\ \forall\, i\ \alpha_i = 0$$

1.1.0.1 Convex combinations

of vectors are linear combinations $\sum \alpha_i \vec{\mathbf{v_i}}$ s.t $\alpha_i \geq 0\ \forall\ i$ and $\sum \alpha_i = 1$

1.1.0.2 Convex Hull $conv(A)$

of $\{\vec{\mathbf{v_i}}\}$ is the set of all convex combinations of the vectors in $\{\vec{\mathbf{v_i}}\}$
$conv(\sigma(A)) = \bigcap_{\{F\}} W(F^{-1}AF)$ with $\{F\}$ being all nonsingular changes of bases.

The

1.1.0.3 Span

of a set of vectors $\mathcal{L}\{S\}$ is the set of all linear combinations of S

A collection of vectors S is a

1.1.0.4 Spanning Set

for V if $\mathcal{L}\{S\} = V$

1.2 Linear Subspaces

W is a subspace of a vector space V if

- W contains the 0 vector
- W is closed under linear combinations. ($x, y \in W$, $a, b \in \mathbb{C} \implies ax + by \in W$)
- For any V, the smallest possible subspace is $\{0\}$ and the larget possible subspace is V itself
- For any $\{\vec{\mathbf{v}}_i\}$, $span(\vec{\mathbf{v}}_i)$ is a subset
- A subspaceof W is **invariant under a maping A** if $AW \subset W$. A will map vectors in W to W.

1.3 Dimension

A vector space V has dimension n if $\exists\, n$ linearly independent elements of V.
If $\{v_1, ...v_n\}$ are linearly independent, then $dim\left(\{v_1, ...v_n\}\right) = n$
If $dim(\{v_1, ...v_n\}) < n$, then the homogenous equation $\sum_{i=1}^{n} \alpha_i \vec{\mathbf{v}}_i$ always has a nontrivial solution

1.4 Bases

A basis of a vector space V is a subset $S \in V$ which contains vectors that are linearly independent and span V
In other words, a basis is a linearly independent spanning set of V

Pierson Guthrey
pguthrey@iastate.edu

- S is a basis of V if and only if $span\{S\} = V$ and each $\vec{\mathbf{v}} \in V$ can be expressed by at least one combination of elements of S
- S is a basis of V if and only if the elements of S are linearly independent and each $\vec{\mathbf{v}} \in V$ can be expressed by by at most one combination of elements of S
- Every vector space has a basis
- The number of elements in the basis is unique
- If $span(\{v_i\})$ is linearly independent, this implies $\forall\, w \in V \;\; \exists\,!\, \alpha_i$ such that $\vec{\mathbf{w}} = \sum_{i=0}^{n} \alpha_i \vec{\mathbf{v}}$

1.4.0.5 Hamel Basis We say $S \subset V$ is a Hamel Basis if $\forall\, x \in V$ there exists a unique set of scalars $\{c_1, ..., c_n\}$ and vectors in S $\{\vec{\mathbf{s}}_1, ..., \vec{\mathbf{s}}_n\}$ such that

$$x = \sum_{i=1}^{n} c_n \vec{\mathbf{s}}_n$$

The Hamel Basis is the default basis

1.5 Norms $\|\cdot\|$

A norm on a linear space X is a function $\|\cdot\| : X \to \mathbb{R}$ with the properties:

- $\|x\| \geq 0 \;\forall\, x \in X$ (nonnegative)
- $\|\lambda x\| = |\lambda| \, \|x\| \;\; \forall\, x \in X, \lambda \in \mathbb{C}$(homogeneous)
- $\|x + y\| \leq \|x\| + \|y\| \;\; \forall\, x, y \in X$ (triangle inequality)
- $\|x\| = 0 \iff x = 0$ (stricly positive)

1.5.1 Normed Linear Space

A normed linear space $(X, \|\cdot\|)$ is a linear space X equipped with a norm $\|\cdot\|$ and metric $d(x, y) = \|x - y\|$ (or $d(x, y) = f(\|x - y\|)$).

Normed linear spaces have the properties:

- $d(x + z, y + z) = d(x, y) \;\forall\, x, y, z \in X$ (translational invariance)
- $d(\lambda x, \lambda y) = |\lambda| \, d(x, y) \;\forall\, \lambda \in \mathbb{C}$ (homogeneity)

Example: $d(x, y) = \frac{\|x-y\|}{\|x-y\| + 1}$

1.6 Vector Norms $\|\cdot\|_p$

$$\|x\|_p = \left(\sum |x_i|^p \right)^{\frac{1}{p}}$$

1.6.0.1 Sum Norm, $p = 1$

$$\|x\|_{sum} = \|x\|_1 = \sum |x_i|$$

1.6.0.2 Euclidean Norm, $p = 2$

$$\|x\|_{Euclidean} = \|x\|_2 = \sqrt{\sum |x_i^2|}$$

1.6.0.3 Maximum Norm, $p = \infty$

$$\|x\|_{max} = \|x\|_\infty = max \{|x_i|\}$$

All finite dimensional norms behave similarly with respect to convergence. Convergence with respect to infinite dimensional spaces depend highly on the norm.

Given a vector x in $\mathbb{C}^N$,

$$\|x\|_{p+a} \leq \|x\|_p \ \forall \ p \geq 1 \ \forall \ a \geq 0$$

$$\|x\|_p \leq \|x\|_r \leq n^{\left(\frac{1}{r} - \frac{1}{p}\right)} \|x\|_p \ \forall \ p \geq 1 \ \forall \ a \geq 0$$

Examples: $\|x\|_2 \leq \|x\|_1$ and $\|x\|_1 \leq \sqrt{n}\|x\|_2$.

2 Linear Mappings

For a linear map (homomorphism) $L : \mathbb{C}^n \to \mathbb{C}^m$ denoted $V \stackrel{L}{\to} W$,
Range $R(L) \subseteq W = \{Lv; v \in V\}$
Nullspace $N(L) \subseteq V = \{v; Lv = 0\}$
and $dim(N(L)) + dim(R(L)) = dim(V)$

Linearity for a map L from a vector space V such that $(v, u) \in V$, $(\alpha, \beta) \in \mathbb{R}$ means

$$L(\alpha v + \beta u) = \alpha L(v) + \beta L(u)$$

If V has basis $\{\vec{\mathbf{e}}_i\}$ and W has basis $\left\{\vec{\mathbf{f}}_i\right\}$ then for any $\vec{\mathbf{v}} \in V$,

$$\vec{\mathbf{v}} = \sum v_i \vec{\mathbf{e}}_i$$
$$L\vec{\mathbf{v}} = \sum v_i L\vec{\mathbf{e}}_i = \sum_i v_i \sum_j L_{ji} \vec{\mathbf{f}}_i$$

A linear map $V \stackrel{L}{\to} W$ is represented by

$$\vec{\mathbf{w}} = L\vec{\mathbf{v}}$$
$$\begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = \begin{pmatrix} L_{11} & \dots & L_{1n} \\ \vdots & \ddots & \vdots \\ L_{m1} & \dots & L_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Pierson Guthrey
pguthrey@iastate.edu

2.0.0.4 Composition For a mapping $V \overset{L}{\to} W \overset{M}{\to} X$ with $dim(n), dim(m), dim(p)$ respectively, the mapping $V \overset{N}{\to} X$ is given by

$$N_{p\times n} = M_{p\times m} \circ L_{m\times n}$$
$$or$$
$$n_{ij} = \sum_k m_{ik} l_{kj}$$

2.0.0.5 Conjugation To change the basis of a vector space, you must also change the linear maps.
Given $V \overset{L_e}{\to} V$ in basis $\{\vec{\mathbf{e}}_1, ..., \vec{\mathbf{e_n}}\}$
and $V \overset{L_f}{\to} V$ in basis $\left\{\vec{\mathbf{f}}_1, ..., \vec{\mathbf{f_n}}\right\}$
related by $\vec{\mathbf{v}}_e = F\vec{\mathbf{v}}_f$, as in F is the basis vectors of the new coordinate system expressed in terms of the old basis vectors.

$$\begin{aligned} \vec{\mathbf{w}}_e &= L_e\vec{\mathbf{v}}_e \\ F\vec{\mathbf{w}}_f &= L_e F\vec{\mathbf{v}}_f \\ \vec{\mathbf{w}}_f &= (F^{-1}L_e F)\vec{\mathbf{v}}_f \end{aligned}$$
$$L_f = F^{-1}L_e F$$

And L_f is known as a conjugate of L_e.

3 Matrices

For convention, a matrix A has n columns and m rows.
$\mathbb{C}^{m\times n}$ is the space of $m \times n$ matrices, which is a vector space of dimension mn
If $m = n$, then you have a ring with unit I (the identity matrix s.t. $I\vec{\mathbf{v}} = \vec{\mathbf{v}}\ \forall\ v$), although there is not generally an inverse.

3.0.0.6 Matrix Transpose Complex Conjugate Transpose (also known as the adjoint) $A* \in \mathbb{C}^{m\times n}$ of $A \in \mathbb{C}^{m\times n}$ is the matrix s.t $a_{ij} = \overline{a_{ji}}$
Defined by $\langle \vec{\mathbf{v}}, A\vec{\mathbf{w}}\rangle = \langle A^*\vec{\mathbf{v}}, \vec{\mathbf{w}}\rangle \qquad \forall\ v, w$

- The adjoint is unique
- Also known as the Hermitian transpose
- $tr(A^*) = \overline{tr(A)}$
- $det(A^*) = \overline{det(A)}$

The **real transpose** is defined as $(A^T)_{ij} = A_{ji}$

- $tr(A^T) = tr(A)$
- $det(A^T) = det(A)$
- If A is real, then $A^T = A^*$
- $(AB)^T = B^T A^T$

3.0.0.7 Symmetry $A \in \mathbb{C}^{n\times n}$ is Hermitian if $A^* = A^H = A$

- Hermitian matrices have real eigenvalues
- $A \in \mathbb{R}^{n\times n}$ is symmetric if $A^T = A$
- Eigenvectors of A corresponding to different eigenvalues are orthogonal.

$$\text{Given } (\lambda_i, \vec{\mathbf{v}}_i) \text{ and } (\lambda_j, \vec{\mathbf{v}}_j), \lambda_i \neq \lambda_j \implies \vec{\mathbf{v}}_i^T \vec{\mathbf{v}}_j = 0$$

- For every eigenvalue of A, the geometric and algebraic multiplicities are the same. If $A \in \mathbb{C}^{m\times m}$, then there exists exactly m linearly independent eigenvectors which can be chosen to be mutually orthonormal. Therefore the spectral decomposition is

$$A = UDU^{-1} = UDU^*$$

  for U unitary and $D \in^{m\times m}$ diagonal with $D_{ii} = \lambda_i$

- If $A \in \mathbb{C}^{m\times n}$, then $A^*A \in \mathbb{C}^{n\times n}$ is Hermitian with eigenvalues $\lambda \geq 0$.

$$(A^*A)^* = A^*A \implies A^*A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$$

$$\vec{\mathbf{x}}^*A^*A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}^*\vec{\mathbf{x}}$$
$$\lambda = \frac{\vec{\mathbf{x}}^*A^*A\vec{\mathbf{x}}}{\vec{\mathbf{x}}^*\vec{\mathbf{x}}} = \frac{(A\vec{\mathbf{x}})^*(A\vec{\mathbf{x}})}{\vec{\mathbf{x}}^*\vec{\mathbf{x}}} \geq 0$$

$A \in \mathbb{C}^{n\times n}$ is skew-symmetric if $A^* = -A$

3.0.0.8 Block Matrices Blocks of matrices can be treated as individual elements of matrices, as long as dimensions are handled properly.
Example. The matrix equation

$$A_{3\times 4}\vec{\mathbf{v}}_{4\times 1} = \vec{\mathbf{b}}_{1\times 3}$$

is equivalent to

$$\begin{pmatrix} B_{2\times 1} & C_{2\times 2} & D_{2\times 1} \\ E_{1\times 1} & F_{1\times 2} & G_{1\times 1} \end{pmatrix} \begin{pmatrix} H_{1\times 1} \\ I_{2\times 1} \\ J_{1\times 1} \end{pmatrix} = \begin{pmatrix} BH + CI + DJ \\ EH + FI + GJ \end{pmatrix}$$

3.1 Elementary Matrix Forms

3.1.0.9 Diagonal Matrices All of the entries above and below (off-diagonal) the diagonal are 0.

$$D = \begin{pmatrix} x & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & x \end{pmatrix}$$

A **rectangular diagonal** matrix has the form

$$\begin{pmatrix} D \\ 0 \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} D & 0 \end{pmatrix}$$

3.1.0.10 Triangular Matrices All of the entries either above or below the diagonal are 0, known as Lower triangular and Upper triangular respectively. It is **strictly lower/upper triangular** if the diagonals are also all 0.

$$R = \begin{pmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & * \end{pmatrix} \qquad L = \begin{pmatrix} * & \dots & 0 \\ \vdots & \ddots & \vdots \\ * & \dots & * \end{pmatrix}$$

- Triangular matrices are assumed to be square
- The inverse of a nonsingular lower (upper) triangular matrix is a lower (upper) triangular matrix
- The product of two upper (lower) triangular matrices is an upper (lower) triangular matrix, since
- The eigenvalues of a triangular matrix lie on the diagonal, and the eigenvectors are easy to find:

$$(R - \lambda I)\vec{\mathbf{v}}_{\mathbf{i}} = \vec{\mathbf{0}}$$
$$\begin{pmatrix} \lambda_1 - \lambda_i & a_{12} & a_{13} \\ 0 & 0 & a_{23} \\ 0 & 0 & \lambda_n - \lambda_i \end{pmatrix} \begin{pmatrix} ? \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

3.1.0.11 Banded All of the entries above and below a certain 'distance' from the diagonal are 0.

$$\begin{pmatrix} * & * & \dots & 0 \\ * & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & * & * \end{pmatrix}$$

- Matrices with a nonzero first superdiagonal and first subdiagonal are known as **tridiagonal**

3.1.0.12 Permutation Matrix Each row and column has exactly one 1 and all other entries are 0.

$$P_\pi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3.1.0.13 Permutations A permutation $\pi : 1, ..., m \to 1, ..., m$ has the form

$$\begin{pmatrix} 1 & 2 & \dots & 4 \\ \pi(1) & \pi(2) & \dots & \pi(m) \end{pmatrix}$$

and can be represented by a permutation matrix P_π. Note:
$sign(\pi) = \pm 1 = (-1)^n = (-1)^{det(P_\pi)}$, n the number of pairwise interchanges that produce the permutation

3.1.0.13.1 Permutation Similarity Matrix of matrix A is a matrix $P_\pi A P_\pi^T$

3.2 Matrix Properties

3.2.0.14 Matrix Invariants For a given $A \in \mathbb{C}^{n \times n}$,
Let $V_k =$ subset of k elements of $(1, ..., n)$
Let $\{\pi\}$ be the collection of all possible permutations of V

Pierson Guthrey
pguthrey@iastate.edu

Then we define $S_k(A)$

$$S_k(A) = \sum_{V_k} \sum_{\{\pi\}} sign(\pi) \prod a_{i,\pi(i)}$$

All $S_k(A)$ are objective (do not change with coordinate systems)

3.2.0.15 Trace of a Matrix $tr(A) = S_1(A) = a_{11} + a_{22} + \ldots + a_{nn}$
The sum of all diagonal entries.

- $tr(AB) = tr(BA)$
- $tr(A^*A) = \sum\limits_{i=1}^{n} a_j^* a_j = \sum\limits_{i=1}^{n} \|a_j\|_2^2$

A sphere of radius 1 transformed using a symmetric A into an ellipsoid will have $|tr(A)| = \sum semiaxes$

3.2.0.16 Determinant of a Matrix $det(A) = S_3(A) = a_{11}a_{22}\ldots a_{nn} + ...$
Note: $|S_k(A)| = |S_3(k)|$ for $k = 2, 3, 4, 5..$
Facts
$det(AB) = det(A)det(B)$ $\quad$ $det(I) = 1$ $\quad$ $det(A^{-1}) = \frac{1}{det(A)}$ (assuming A^{-1} exists)
$det(A) \neq 0 \iff$ the matrix is nonsingular $\iff$ nullspace is the zero vector
A cube of volume 1 transformed using A into a parallelpiped will have volume $|det(A)|$

3.2.0.17 Matrix Inverse Any nonsingular matrix has a unique inverse A^{-1} s.t $AA^{-1} = A^{-1}A = I$

- If $A, B \in \mathbb{R}^{n\times n}$ are invertible ($\det A \neq 0, \det B \neq 0$) then so it AB and $(AB)^{-1} = B^{-1}A^{-1}$

The following are equivalent for $A \in \mathbb{C}^{n\times n}$

1. A^{-1} exists.
2. $\det(A) \neq= 0$
3. $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a unique solution for any $\vec{\mathbf{b}}$
4. The rows of A are linearly independent
5. The columns of A are linearly independent
6. $N(A) = \left\{\vec{\mathbf{0}}\right\}$ and thus $null(A) = 0$
7. $R(A) = \mathbb{C}^n$ and thus $rank(A) = n$

Column Space: $C(A)$ is the linear combination of all column vectors of a matrix. You can solve $Ax = b$ exactly when b is in the column space of A. If the column vectors span R^m, then $C(A)$ is a subspace of R^m, but may be of a lesser dimension. The column space is not in general a subspace.

$$R(A) = \{\vec{\mathbf{y}} \in \mathbb{C}^n : \vec{\mathbf{y}} = A\vec{\mathbf{x}} \text{ for some } \vec{\mathbf{x}} \in \mathbb{C}^n\}$$

Pierson Guthrey
pguthrey@iastate.edu

- $dim(R(A)) = rank(A)$ is the **rank** of A
- $rank(A) + null(A) = n$

Null Space: $N(A)$ is the space of all solutions to $Ax = 0$. If the row vectors span R^n, then the nullspace is a supspace of R^n. Usually this is found by putting the matrix in echelon form and then finding the space spanned by the special solutions.

$$N(A) = \left\{ \vec{\mathbf{x}} \in \mathbb{C}^n : A\vec{\mathbf{x}} = \vec{\mathbf{0}} \right\}$$

- For any matrix A, $\vec{\mathbf{0}} \in N(A)$
- $dim(N(A)) = null(A)$ is the **nullity** of A

Special solutions: When fidning a null space, set each free variable in turn to 1 and find the vector that solves $rref(A)x = 0$. Each of those is a special solution. The null space is the space spanned by the linear combination of all of the special solutions.

Rank of the Matrix: The $rank(A)$ is... It is always the case that $rank(A) \leq m$ AND $rank(A) \leq n$. A is of full rank iff
Full column rank: If r=n, $null(A) = 0$. You have n pivot variables and no free variables. This means you have at most 1 solutions to $Ax = b$.

- Full row rank: If r=m, you can solve $Ax = b$ for any b. There are an infinite number of solutions. You have m pivot variables and n-m free variables.
- Full rank: There is one unique solution for any b for $Ax = b$.

Solvability Condition: The condition on b such that $Ax = b$ has solutions. Namely, b must be in $C(A)$.

Complete Solution: The complete solution to $Ax = b$ is the linear combination of the particular solution x_p and all of the vectors in $N(a)$. $x_{complete} = x_{particular} + x_{null}$.

Echelon form: A is in echelon form if it has the shapce resulting from Gaussian Elimination. If GA is performed on the rows, it is known as Row Echelon Form. If GA is performed on the columns, it is known as Column Echelon Form.

- All nonzero rows/columns are above any rows/columns of all zeroes.
- The leading coefficient (first nonzero number from the left, aka the pivot) of a nonzero row/column is always strictly to the right of the leading coefficient of the row above it.
- All entries in a cloumn below a leading entry are zeroes.

Reduced row echelon form (aka canoonical form): A matrix is in this form if it meets all the criteria of Echelon Form, but in addtion also satisfies:
• Every leading coefficient is 1 and is the only nonzero entry in it's column.

3.3 Matrix Norms

Matrix norms treat a matrix A as an operator and come in three common varieties: induced norms, Schatten norms, and entrywise norms. Matrix norms have the property $\|AB\| \leq \|A\| \, \|B\|$.

Pierson Guthrey
pguthrey@iastate.edu

3.3.1 Induced Norms

Given a vector norm $\|\cdot\|_p$, let $A \in \mathbb{C}^{n\times m}$, the subordinate norm of A is

$$\|A\|_{p,q} = \sup_{x\neq 0, x\in\mathbb{C}^n} \frac{\|Ax\|_q}{\|x\|_p} = \sup_{\|x\|_p=1, x\in\mathbb{C}^n} \|Ax\|_q$$

If $p = q$, we have the induced norm

$$\|A\|_p = \sup_{x\neq 0, x\in\mathbb{C}^n} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p=1, x\in\mathbb{C}^n} \|Ax\|_p$$

- $p = 1$ leads to the maximum absolute column sum

$$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$$

- $p = 2$ leads to the largest singular value of A, ie the square root of the largest eigenvalue of the positive semi-definite matrix A^*A

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)} = \sigma_{\max}(A)$$

- $p = \infty$ leads to the maximum absolute row sum

$$\|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$

Fun facts

- For any induced norm, $\|A^r\|_p^{r^{-1}} \geq \rho(A)$, where equality holds for $p = 2$.

3.3.2 Schatten Norms

The Schatten norms arise when applying the $p-$norm to the vector of singular values of a matrix. Let $\{\sigma_i\}$ be the singular values of A.

$$\|A\|_p = \left(\sum_{i=1}^{\min(m,n)} \sigma_i^p\right)^{\frac{1}{p}}$$

Schatten norms are unitarily invariant. That is, for any unitary matrices U, V,

$$\|A\|_p = \|UAV\|_p$$

- For $p = 1$, we have the **Nuclear norm** also known as the **Trace norm**

$$\|A\|_* = \sum_{i=1}^{\min(m,n)} \sigma_i$$

- $p = 2$. The Frobenius norm is the finite dimensional analog of the Hilbert Schmidt norm

$$\|A\|_F = \|A\|_2 = \left(\sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|^2\right)^{\frac{1}{2}}$$

Pierson Guthrey
pguthrey@iastate.edu

with the interesting property

$$\|A\|_F = \sqrt{\text{trace}(A^*A)} = \sqrt{\sum_{i=1}^{\min(m,n)} \sigma_i^2}$$

- For $p = \infty$, we have the **spectral norm**

$$\|A\|_* = \max_{1 \leq i \leq \min(m,n)} \sigma_i$$

3.3.3 Entrywise Norms

Entrywise norms treat $A \in \mathbb{C}^{n \times n}$ as a vector of size mn, so they take the form

$$\|A\|_{p,q} = \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}|^p \right)^{\frac{q}{p}} \right)^{q^{-1}}$$

These are not guaranteed to be sub-multiplicative, and therefore are norms that may not be consistent with vector norms.

- In data analysis, a common norm is $L2,1$ norm

$$\|A\|_{2,1} = \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}|^2 \right)^{\frac{1}{2}} \right)$$

- The max norm is not sub-multiplicative and therefore not a true norm

$$\|A\|_{\max} = \|A\|_{\infty,\infty} = \max_{1 \leq i \leq m, 1 \leq j \leq n} |a_{ij}|$$

3.4 Eigenvalues

For a given $A \in \mathbb{C}$, $\lambda \in \mathbb{C}^{n \times n}$ is an eigenvalue of A associated with eigenvector $v \neq 0$ s.t

$$Av = \lambda v$$

and the pair (λ, v) decribe a vector v for which A maps to a a multiple λ of itself.
Eigenvalues of $A^* =$ complex conjugates of eigenvalues of A
Eigenvalues of $A^T =$ eigenvalues of A
$\sum eigenvalues = tr(A)$
$\prod eigenvalues = det(A)$
Eigenvalues are preserved under change of basis (eigenvalues of $A =$ eigenvalues of $F^{-1}AF$) but eigenvectors $\vec{\mathbf{v}}$ will be different.

$$\det(A - \lambda I) = \det(F^{-1}BF - F^{-1}\lambda IF) = \det(B - \lambda I), \vec{\mathbf{v}}' = F\vec{\mathbf{v}}$$

Eigenvalues depend continuously on the matrix entries.

3.4.0.1 Gershgorin Theorem

Let $\mathbb{D}(A) = \bigcup_{i=1}^{n} D_i$ with circles $D_i = \{z \in \mathbb{C} | \, |z - a_{ii}| \leq r_i\}$ with the row sum $r_i = \sum_{i \neq j} = |a_{ij}|$
All eigenvalues are in $\mathbb{D}(A)$
If there are k circles D_i which intersect but are separated from other circles then there are exactly k eigenvalues in the union of the k circles. (Note the eigenvalues are in the union of the circles, not the circles themselves)
If $r_i = 0 \; \forall \; i$, then D_i is the set of eigenvalues in circles of radius 0.
A^T creates different circles (it uses column sums of A)

3.4.0.2 Eigenvectors $\vec{\mathbf{v}}$

Right eigenvectors are $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$
Left eigenvectors are $A^*\vec{\mathbf{w}} = \overline{\lambda}\vec{\mathbf{w}}$ or $\vec{\mathbf{w}}^* A = \lambda\vec{\mathbf{w}}^*$

3.4.0.3 Eigenspace $E(\lambda)$

The subspace of $\mathbb{C}^{n \times n}$ defined by $\{v : Av = \lambda v\}$

3.4.0.4 Characteristic Polynomial $\chi(\lambda)$

$\chi(\lambda) = det(A - \lambda I) = 0$

- Fundamental Theorem of Algebra implies every $n \times n$ matrix A has exactly n eigenvalues, possibly complex, possibly multiple.
- For every eigenvalue there is at least one eigenvector.
- For real A, any complex eigenvalues come in complex conjugate pairs $(\lambda, \vec{\mathbf{v}})$ and $(\overline{\lambda}, \overline{\vec{\mathbf{v}}})$
- **Cayley Hamilton Theorem:** A matrix satisfies its own characteristic polynomial

$$\chi_A(A) = A + a_{n-1}A^{n-1} + ... + a_0 I$$

  Proof is in 507 Notes

3.4.0.5 Multiplicity

- The **Algebraic Multiplicity** of an eigenvalue is its multiplicity as a root of the characteristic polynomial of the matrix. We can see

$$\text{Algebraic multiplicity } \alpha(\lambda) = \dim N_p = \sum_i \dim M_i$$

- The **Geometric Multiplicity** $\gamma(\lambda)$ of an eigenvalue is the number of linearly independent eigenvectors associated with that eigenvalue
- $1 \leq \gamma(\lambda) = \dim M_1 \leq \alpha(\lambda) = \dim N_p$

3.4.0.6 Spectrum

Spectrum $\sigma(A) = \{$eigenvalues of A$\}$
Spectral radius $\rho(A) = max(|\lambda|)$
$\rho(A^*) = \rho(A^T) = \rho(A)$

3.4.0.7 Orthonormal A collection of vectors $\{\vec{\mathbf{v_i}}\}$ is orthonormal if $\langle v_i, v_j\rangle = \delta_{ij}$

3.5 Numerical Range and Radius

3.5.0.8 Numerical Range $W(A) = \{U^*AU : \|U\| \ = 1\}$ for $A \in \mathbb{C}^{n\times n}$
$\sigma(A) \subset W(A)$
$W(\alpha A + \beta I) = \alpha W(A) + \beta$
$W(A \oplus B) \subset W(A) + W(B)$
$W(A^*) = \overline{W(A)}$
$W(A)$ is compact and convex
$conv(\sigma(A)) \subset W(A)$
If A is Hermitian, $W(A) = [\lambda_{\text{min}}, \lambda_{\text{max}}]$

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \implies W(A) = conv(W(A_1)\bigcup W(A_2))$$

$W(A) = \{0\} \implies A = 0$

3.5.0.9 Numerical Radius $r(A) = max(\{|z|, z \in W(A)\})$
$\rho(A) \leq r(A)$
For A normal. $\rho(A) = r(A)$
$r(A^*) = r(A^T) = r(A)$

3.6 Minimal Polynomial

The minimal polynomial $\mu_A(\lambda)$ is the polynomial of lowest degree for which $\mu(A) = 0$

- $\mu_A(\lambda)$ always exists since $\chi_A(A) = 0$
- $\mu_A(\lambda)$ always has the form $\prod\limits_i(\lambda - \lambda_i)^{k_i}$ with $0 \leq k_i \leq \alpha(\lambda_i)$ for eigenvalues $\{\lambda_i\}$ of A
- Relation to Jordan chains: $\mu_A(\lambda) = \prod\limits_i(\lambda - \lambda_i)^{\mu_i}$, where μ_i is the length of the longest chain

3.7 Special Types of Matrices

3.7.0.10 Orthogonal Matrices $Q \in \mathbb{R}^{n\times n}$ is orthogonal if $Q^T = Q^{-1}$

- Q is made up of mutually orthonormal vectors $\vec{\mathbf{q}}_i^T\vec{\mathbf{q}}_j = \delta_{ij}$
- $Q^TQ = I$

3.7.0.11 Unitary $U \in \mathbb{C}^{n\times n}$ is unitary if $U^*U = I \iff U^* = U^{-1}$
The column vectors are orthonormal.

$$\begin{pmatrix} \vec{\mathbf{U}}_1^* \\ \vec{\mathbf{U}}_2^* \\ \vec{\mathbf{U}}_3^* \end{pmatrix}\begin{pmatrix} \vec{\mathbf{U}}_1 & \vec{\mathbf{U}}_2 & \vec{\mathbf{U}}_3 \end{pmatrix} = I$$

Implying $\langle U_i^*, U_j \rangle = \delta_{ij}$
$det(U) = \pm 1 \implies |det(U)| = 1$
$\|U\vec{\mathbf{v}}\| = \|\vec{\mathbf{v}}\| \qquad \forall\, \vec{\mathbf{v}}$
$\langle U\vec{\mathbf{v}}, U\vec{\mathbf{w}} \rangle = \langle \vec{\mathbf{v}}, \vec{\mathbf{w}} \rangle$
If U is real it is called ortogonal and is a rotation matrix
Eigenvalues of U are on the unit circle.

3.7.0.12 Rotation Matrix

A real unitary matrix $U \in \mathbb{R}^{n\times n}$ has one of the following forms:

$$U = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} cos(\theta) & sin(\theta) \\ sin(\theta) & -cos(\theta) \end{pmatrix}$$

$det(U) = 1$ if U is a rotation
$det(U) = -1$ if U is a rotation and reflection

3.7.0.13 Normal Matrix

A square matrix $N \in \mathbb{C}^{n\times n}$ is normal if $N^*N = NN^*$
$W(N) = conv(\sigma(A))$
$\rho(A) = r(A)$

3.7.0.14 Nilpotent Matrix

A is **nilpotent of order** p if

$$A^p = 0 \text{ and } A^{p-1} \neq 0 \qquad \text{for some } p \geq 1$$

As in, p is the lowest power for which A^p becomes the zero matrix

- A is nilpotent if and only if all eigenvalues of A are zero

3.7.0.15 Projection Matrix

A matrix P projects a vector $v \in V$ onto a subspace of V, say U

- $P^2 = P$, and all eigenvalues of P are 0 or 1
- The target subspace U is invariant under P. That is, P is the identity operator of U
- P is idempotent
- Say K is the kernel of W. We have a direct sum $W = K \oplus U$ and every vector in W can be decomposed as $x = u + k$ with $u = Px$ and $k = x - Px$.
- The range and kenrel of P are complementary

$$\mathbb{C}^m = range(P) \oplus null(P), range(P) \cap null(P) = \left\{\vec{\mathbf{0}}\right\}$$

- $Q = I - P$ is also a projector. The kernel and range of P are the range and kernel of Q and vice-versa.

$$range(P) = null(I-P), range(I-P) = null(P), null(P) \cap null(I-P) = \left\{\vec{\mathbf{0}}\right\}$$

- In an infite dimensional vector space, the spectrum is contained in $\{0,1\}$. The only eigenvalues of a projection are 0 (the eigenspace of which is the kernel of P) or 1 (the eigenspace of which is the range of P).

Pierson Guthrey
pguthrey@iastate.edu

3.8 Matrix Factorizations

3.8.1 LU Factorization

$$PA = LU$$

or

$$PA = LDU$$

- L is lower triangular
- D is diagonal
- R is upper triangular
- P is a permutation matrix to ensure the first element of L is nonzero
- All square matrices can be factored this way
- If A is invertible, it admits an LDU factorization if and only if all leading principle minors are nonzero

3.8.1.1 Cholesky Decomposition

If A is Hermitian (or symmetric) and positive definite, for a LU or LDU factorization we see that $U = L^*$ so

$$A = LL^*$$

or

$$A = LDL^*$$

- Always exists and is always unique
- Numerically stable compared to other LU factorizations

3.8.2 QR Decomposition

$$A = QR$$

For an orthogonal matrix Q and an upper triangular matrix R. If A is invertible, then the requiring the diagonal entries of R to be positive makes the decomposition unqiue.
If A is not full rank,

$$A_{m\times n} = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}_{m\times m} \begin{pmatrix} R_1 \\ 0 \end{pmatrix}_{m\times n} = Q_1 R_1$$

If A has n linearly independent columns,

- Q_1 is $m \times n$ and forms an orthonormal basis for the column space of A
- Q_2 is $m \times (m - n)$

3.8.3 Schur Factorization

If $A \in \mathbb{C}^{n\times n}$

$$A = QUQ^{-1} = QUQ^*$$

- Q is unitary, $Q^* = Q^{-1}$
- U is upper triangular and is known as the Schur Form of A
- U is similiar to A and has the eigenvalues of A along its diagonal

Facts

- If A is Hermitian Schur Factorization is identical to Spectral Decomposition
- Every square matrix is unitarily similar to an upper triangular matrix (has a Schur Factorization)

3.8.4 Spectral Decomposition

If $A \in \mathbb{C}^{n\times n}$ is diagonalizeable (distinct eigenvectors), then

$$A = V\Lambda V^{-1}$$

- Λ is diagonal formed from the eigenvalues of A
- V are the corresponding eigenvectors of A

Facts

- Also known as Eigendecomposition
- Existence: Since $AV = \Lambda V$, V^-1 exists if the eigenvectors are distinct
- Requiring that each eigenvalue's geometric multiplicity be equal to its algebraic multiplicity is necessary and sufficient
- If A is also symmetric, and we normalize so $V^TV = I$, then $A = V\Lambda V^T$
- If A is Hermitian Spectral Decomposition is identical to Schur Factorization

3.8.5 Unitary Diagonalizeability

If A is normal then

$$A = QDQ^*$$

for some unitary matrix Q.

- A is unitarily diagonalizeable if and only if A is normal.

3.8.6 Rank Factorization

Any matrix $A \in \mathbb{C}^{m\times n}$ that has rank r, can be factored:

$$A = CF$$

- $C_{m\times r}$ has columns that form a basis for the column space of A
- $F_{r\times n}$ has the coefficients for each row of A in terms of the basis C. That is,

$$a_j = f_{1j}c_1 + f_{2j}c_2 + ... + f_{rj}c_r$$

- There is no restriction on A

3.8.7 Polar Decomposition

For a square matrix A,

$$A = QH$$

- $Q = UV^*$ is unitary
- $H = V\Sigma V^*$ is Hermitian and positive definite
- H is unique

3.8.8 Singular Value Decomposition

For any matrix $A_{m\times n}$,

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^*_{n\times n}$$

- U is a unitary matrix whose columns are the **left singular vectors** of A. The left singular vectors of A are the eigenvectors of AA^*
- Σ is a retangular diagonal matrix with the **singular values** σ_i of A along the diagonal. The singular values of A are the nonzero eigenvalues of AA^* and A^*A. The convention is to write them in nonincreasing order, and so $\sigma_1 \geq \sigma_2 \geq ....$
- V^* is a unitary matrix where the columns of V are the **right singular vectors** of A. The right singular vectors of A are the eigenvectors of A^*A
- Σ is unique, but U,V^* are not
- The SVD exists for all matrices $A \in \mathbb{C}^{n\times m}$. (Proof by construction)

If $m \geq n$, let $p = \min(m, n)$

$$U\Sigma V^* = \begin{pmatrix} \vec{\mathbf{u}}_1 & \cdots & \vec{\mathbf{u}}_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_p \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_n \end{pmatrix}^*$$

If $m \geq n$, let $p = \min(m, n)$

$$U\Sigma V^* = \begin{pmatrix} \vec{\mathbf{u}}_1 & \cdots & \vec{\mathbf{u}}_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & & 0 & \dots \\ & \ddots & & 0 & \dots \\ & & \sigma_p & 0 & \dots \end{pmatrix} \begin{pmatrix} \vec{\mathbf{v}}_1 & \cdots & \vec{\mathbf{v}}_n \end{pmatrix}^*$$

- Geometric interpretation: The image of the unit sphere in $\mathbb{R}^N$ under the map $A = U\Sigma V^*$ is a hyperellipse in $\mathbb{R}^M$. Take $x \in \mathbb{R}^N$ with $\|x\|_2 = 1$. Notice $\|V^*x\|_2 = 1$. V^* rotates and or reflects x. $D(V^*x)$ is a stretching and or contracting of certain elements of V^*x. U rotates and or reflects DV^*x.
  If we are in $\mathbb{R}^2$, this ellipse has semimajor axis σ_1 and semiminor axis σ_2.
- The SVD can be used to rewrite A as the sum of r rank one matrices

  $$A_r = \sigma_1\vec{\mathbf{u}}_1\vec{\mathbf{v}}_1^* + ... + \sigma_r\vec{\mathbf{u}}_r\vec{\mathbf{v}}_r^*$$

  With the interesting property that

  $$\|A - A_k\|_2 = \begin{cases} \sigma_{k+1} & k = 1, 2, ..., r-1 \\ 0 & k = r \end{cases}$$

Which solves the optimization problem of finding the closest rank k matrix to A

$$\|A - A_k\|_2 = \inf\left\{\|A - B\|_2 : B \in \mathbb{C}^{m\times n}, rank(B) = k\right\}$$

3.8.9 Moore Penrose Pseudoinverse OR Generalized Inverse

$A^\dagger$ is the generalization of the matrix inverse that satisfies

$$A^\dagger = \left(A^* A\right)^{-1} A$$

or using the SVD $A = U\Sigma V^*$:

$$A^\dagger = V\Sigma_1^{-1}U^*$$

Where $\Sigma_1^{-1} = \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ is th einverse othe nonzero block of Σ

- $\left(A^\dagger\right)^\dagger$
- $AA^\dagger A = A$, where $AA^\dagger$ maps columns to themselves but need not be the identity
- $A^\dagger AA^\dagger = A^\dagger$, where $A^\dagger$ is the weak inverse
- $\left(AA^\dagger\right)^* = AA^\dagger$, therefore $AA^\dagger$ is Hermitian
- $\left(A^\dagger A\right)^* = A^\dagger A$, therefore $A^\dagger A$ is Hermitian

Facts about the pseudoinverse

- $A^\dagger$ exists for any matrix
- If $A \in \mathbb{R}^{n\times n}$, then $A^\dagger \in \mathbb{R}^{n\times n}$
- If A is invertible, $A^\dagger = A^{-1}$
- $\left(A^T\right)^\dagger = \left(A^\dagger\right)^T$
- $\overline{A}^\dagger = \overline{A^\dagger}$
- $(A^*)^\dagger = \left(A^\dagger\right)^*$
- $(\alpha A)^\dagger = \alpha^{-1}A^\dagger$ if $\alpha \neq 0$

3.9 Matrix Nullspace $N(A)$

For a matrix A, the nullspace is the set of all vectors $N(A)$ such that if $v \in N(A)$, then $Av = 0$. That is,

$$x \in N(A) \iff Ax = 0$$

- $N(A^{k+1}) \supseteq N(A^k)$ since $x \in N(A^k) \iff A^k x = 0$
- If $A^k = A^{k+1}$ for some k, then $A^k = A^{k+1} = A^{k+2} = ... = A^m \; \forall \; m \geq k$

If A is nilpotent of order p, then

$$\{0\} \subsetneqq N(A) \subsetneqq N(A^2) \subsetneqq ... \subsetneqq N(A^p) = N(A^{p+1}) = ...$$

denoted $N_0, N_1, N_2, ..., N_p, N_{p+1}, ...$ respectively

3.9.0.1 Generalized Nullspace $\mathcal{N}(A) \quad = N_p$

If A is not nilpotent, then $\mathcal{N}(A) = \{0\}$
Define M_k as the complement of N_{k-1} in N_k, so

$$N_p = N_{p-1} \oplus M_p = N_{p-2} \oplus M_{p-1} \oplus M_p$$

and

$$N_1 = N_0 \oplus W_1 = W_1$$

so we see

$$C^n = M_1 \oplus M_2 \oplus ... \oplus M_p$$

3.9.0.2 Generalized Eigenvectors

v is a Generalized Eigenvector of order k for eigenvalue $\lambda = 0$ if

$$v \in M_k \iff A^k v = 0 \text{ where } A^{k-1} \neq 0$$

- M_1 are the eigenvectors of A
- A maps $M_k \to M_{k-1}$ for $k \geq 2$ and is one-to-one
- $\dim(M_k) \leq \dim(M_{k-1})$

3.10 Jordan Normal Form

The Jordan Normal form generalizes the eigendecomposition to cases where there are repeated eigenvalues and the matrix cannot be diagonalized. The Jordan Normal Form for a matrix A is

$$A = VJV^{-1}$$

where V is a basis formed from the generalized eigenvectors of A. J has the form

$$J_{?\times ?} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_? \end{pmatrix} \qquad \text{with blocks} \qquad J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}_{?\times ?}$$

- J is upper triangular
- VJV^{-1} is not unique since we can rearrange the blocks and corresponding eigenvectors
- Each block corresponds to one eigenvector
- The number of blocks to a given λ_i is equal to $\gamma(\lambda_i)$
- The size of a block is the length of the chain
- A is diagonalizeable if and only if all blocks have size $|x|$
- For $A \in \mathbb{C}^{n\times n}$, if $n \leq i$ for some small unkown i, the chain structure is uniquely determined by α, γ, μ

3.10.0.3 Example If $A \in \mathbb{C}^{4\times 4}$ with a single eigenvalue λ, we consider all possibilities (permutations are equivalent)
Note that μ is the length of the longest chain

		$\gamma(\lambda)$	$\alpha(\lambda)$	μ
single chain	(4)	1	4	4
two chains	$(3,1)$	2	4	3
	$(2,2)$	2	4	2
	$(2,1,1)$	3	4	2
	$(1,1,1,1)$	4	4	1

A Jordan Chain is the column of eigenvectors in a table of the following form:
Say $\dim M_1 = 4$, $\dim M_2 = 3$, and $\dim M_3 = 1$, then

$$\begin{array}{c|c|c|c|cc}
M_3 & v_1^{(3)} & & & & = N_3 \setminus N_2 \\
M_2 & v_1^{(2)} = Av_1^{(3)} & v_2^{(2)} & v_3^{(2)} & & = N_2 \setminus N_1 \\
M_1 & v_1^{(1)} = Av_1^{(2)} & v_2^{(1)} = Av_2^{(2)} & v_3^{(1)} = Av_3^{(2)} & v_4^{(1)} & = N_1 \setminus N_0 \\
 & \{0\} & & & & = N_0
\end{array}$$

- Each
$$v_i^{(k)} = \begin{cases} i^{th} \text{ k-dimensional generalized eigenvector} & k > 1 \\ i^{th} \text{ k-dimensional eigenvector of } A & k = 1 \end{cases}$$

- The Jordan Chains are subspaces which are invariant under A?

- If A is nilpotent we can choose a basis v and see
$$\begin{aligned} v &= \begin{pmatrix} v_1^{(1)} & v_1^{(2)} & v_1^{(3)} & v_2^{(1)} & v_2^{(2)} & v_3^{(1)} & v_3^{(2)} & v_4^{(1)} \end{pmatrix} \\ Av &= \begin{pmatrix} 0 & v_1^{(1)} & v_1^{(2)} & 0 & v_2^{(1)} & 0 & v_3^{(1)} & 0 \end{pmatrix} \end{aligned}$$

3.11 Least Squares Problems

We want to solve $Ax = b$ for a $m \times n$ matrix of rank r.
The least squares solution $\hat{x}$ to $Ax = b$ is the x with smallest $\|\cdot\|_2$ among all x that minimize $\|Ax - b\|_2$ In general, this is
$$\hat{x} = A^\dagger b = V\Sigma^\dagger U^* b$$

In MatLab, $x = A \setminus b$ gives $\hat{x}$ using QR

3.11.1 Exactly Solvable Case: $m = n = r$

If A is square, nonsingular ($m = n = r$),

- $Ax = b$ has a solution if and only if $b \in R(A)$
- Given a solution x, and $n \in N(A)$, $x + n$ is also a solution since $A(x + n) = Ax = b$
- Solution is unique if and only if $N(A) = \{0\}$

Pierson Guthrey
pguthrey@iastate.edu

3.11.1.1 Solution Methods

- The soluton is
$$x = A^{-1}b$$
But this is not practical to compute, but proves the existence of a solution
- Using the LU decomposition
$$Ax = P^{-1}LUx = b \qquad \begin{cases} A = LU \\ Ly = Pb \\ Ux = y \end{cases}$$
Which can be solved with forward/backward substitution.
- Using QR (thrice the effort of LU)
$$Ax = QRP^{-1}x = b \qquad \begin{cases} A = QRP^{-1} & \\ Qz = b & z = Q^*b \\ Ry = z & \text{triangular} \\ P^{-1}x = y & x = Py \end{cases}$$
- Using SVD (requires the most effort)
$$Ax = U\Sigma V^*x = b \qquad \begin{cases} A = U\Sigma V^* & \\ Uz = b & z = U^*b \\ \Sigma y = z & \text{diagonal} \\ V^*x = y & x = Vy \end{cases}$$

3.11.2 Overdetermined case: $m > n$, $n = r$

N equations with 2 unknowns:
$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

- $N(A) = \{0\}$
- $R(A)$ is not the whole space (is a proper subset of whole space), which implies a solution need not exist
- The **residual** $r(x) = r = Ax - b$ is a function of x

3.11.2.1 Approximate solution method: Minimize the residual using normal equations

You can minimize with respect to any norm but $\|\cdot\|_2$ is easiest. If we define $\hat{b}$ as the orthogonal projection of b onto $R(A)$,
$$r(x)^2 = \hat{b}^2 + \left(\hat{b} - Ax\right)^2 \quad \text{I don't think this is right}$$
so we can make $b - Ax = 0$
$$r \perp R(A) \iff r \in N(A^*)^\perp \iff A^*r = 0$$
But
$$r \in N(A^*)^\perp \implies A^*(Ax - b) = 0 \implies A^*Ax = A^*b$$

Pierson Guthrey
pguthrey@iastate.edu

- $A^*Ax = A^*b$ is the **Normal Equation** for $Ax = b$, which minimizes the sum of square differences between the left and right sides
- Since $K(A^*A) \leq K(A)^2$, the condition number for the normal equations could be very large

so

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = x = \left(A^*A\right)^{-1} A^*b$$

3.11.2.2 Approximate solution method: Minimize the residual using QR

$$\|r\|_2 = \|Ax - b\|_2 = \|QRx - b\|_2 = \|Rx - Q^*b\|_2$$

In case 2, if we define $c = Q^*b$

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} \implies Rx - c = \begin{pmatrix} R_1 \\ 0 \end{pmatrix} x - \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} R_1x - c_1 \\ -c_2 \end{pmatrix}$$

Find x such that $R_1x - c_1 = 0$ so

$$\|r\|_2^2 = \left\| \begin{pmatrix} R_1x - c_1 \\ -c_2 \end{pmatrix} \right\|_2^2 = \|R_1x - c_1\|_2^2 + \|c_2\|_2^2 = \|c_2\|_2^2$$

Which can be found via backwards substitution on

$$x = R_1^{-1}c_1$$

3.11.3 Underdetermined case: $m < n$, $m = r$

3.11.3.1 Solution Method: SVD

Define $c = U^*b$ and $y = V^*x$, so

$$\|r\|_2^2 = \|\Sigma y - c\|_2$$

But since $\Sigma y = \begin{pmatrix} \Sigma_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \Sigma_1 y_1$, we set $y_1 = \Sigma_1^{-1}c$ and let y_2 be such that $\|y_2\|_2 = 0$ so that

$$\|x\|_2^2 = \|Vy\|_2^2 = \|y\|_2^2 = \|y_1\|_2^2 + \|y_2\|_2^2 = \left\|\Sigma_1^{-1}c\right\|_2^2$$

3.11.3.2 Solution Method: QR

Define $c = Q^\dagger b$, so

$$\|r\|_2^2 = \left\|Rx - Q^\dagger b\right\|_2$$

But since $Rx = \begin{pmatrix} R_1 & R_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = R_1x_1 + R_2x_2$, with x_2 arbitrary since R_1 is upper triangular but R_2 is arbitrary. We set $x_1 = R_1^{-1}\left(c - R_2x_2\right)$

$$\|x\|_2^2 = \|x_1\|_2^2 + \|x_2\|_2^2 = \left\|R_1^{-1}\left(c - R_2x_2\right)\right\|_2^2 + \|x_2\|_2^2$$

It is unknown how to minimize this.

3.12 Arbitrary m, n, r

Using QR,

$$Ax = QRP^{-1}x = b$$

We define $c = Q^*b$ and $y = P^{-1}x$ and we want to solve $Ry = c$

$$\begin{pmatrix} R_1 & R_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

With R_1 upper triangular and R_2 arbitrary

$$\|x\|_2^2 = \|R_1y_1 + R_2y_2 - c_1\|_2^2 + \|c_2\|_2^2$$

So we set $y_1 = R_1^{-1}(c_1 - R_2y_2)$ and with y_2 arbitrary

3.13 Linear Programming

Minimize or maximize $f_0(x)$ subject to

$$\begin{cases} f_1(x) \leq 0 \\ \quad\vdots \\ f_m(x) \leq 0 \\ Ax - b = 0 \end{cases} \qquad \text{or} \qquad \begin{cases} f_1(x) \geq 0 \\ \quad\vdots \\ f_m(x) \geq 0 \\ Ax - b = 0 \end{cases}$$

Solve using Simplex method

4 Nonnegative and Stochastic Matrices

A matrix $M = [m_{ij}] \in \mathbb{R}^{n\times n}$ is defined as one of the following if the given condition holds for all elements m_{ij}

	condition	denoted
nonnegative	$m_{ij} \geq 0$	$M > 0$
positive	$m_{ij} > 0$	$M \geq 0$
semipositive	$M \geq 0$ and $M \neq 0$	$M \gneq 0$

- The absolute value of $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ is the matrix $|A| = [|a_{ij}|]$
- The **dominant eigenvalue** of A is the eigenvalue such that $|\lambda| = \rho(A)$ (but with eval stuff)

4.0.1 Facts

Let $A, B, M, N, P, N_1, M_1 \in \mathbb{C}^{n\times n}$ and let $M \geq N \gneq 0$, let $P > 0$, let $0 \leq N_1 \leq M_1$ and let $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \mathbb{C}^n$

- $|AB| \leq |A|\,|B|$
- $|A\vec{\mathbf{v}}| \leq |A|\,|\vec{\mathbf{v}}|$
- $NP > 0$ if and only if N does not have a row of $0's$
- $PN > 0$ if and only if N does not have a column of $0's$
- If $w > 0$ then $N\vec{\mathbf{w}} \gneq 0$

- If $w \gneq 0$ then $P\vec{\mathbf{w}} > 0$
- $NN_1 \leq MM_1$
- $\|N_1\|_1 \leq \|M_1\|_1$ and $\|N\|_\infty \leq \|M\|_\infty$
- $\||A|\|_1 = \|A\|_1$ and $\||A|\|_\infty \leq \|A\|_\infty$
- Put with norms: $\lim\limits_{k\to\infty} \left\|A^k\right\|^{\frac{1}{k}} = \rho(A)$
  Corollary: If for all $k \in \mathbb{N}$, $\left|A^k\right| \leq \left|B^k\right|$ then $\rho(A) \leq \rho(B)$
- For any $A \in \mathbb{C}^{n\times n}$ $\rho(A) \leq \rho(|A|)$
- If $0 \leq N \leq M$ then $\rho(N) \leq \rho(M)$
- If $M \geq 0$ then $\rho(M) >\geq 0\rho(M[S])$ where $M[S]$ is the principal submatrix of M determined by S. In particular, $m_{ii} \leq \rho(M)$ for $i = 1, ..., n$
- If $P > N \geq 0$, then $\rho(P) > \rho(N)$ and $\rho(P) > 0$
- For a nonnegative matrix $M \in \mathbb{R}^{n\times n}$, the ith row sum and jth column sum of M are denoted

  $$r_i(M) = \sum_{j=1}^n m_{ij} \qquad \text{and} \qquad c_j(M) = \sum_{i=1}^n m_{ij}$$

  Then

  $$\min_{1\leq i\leq n} r_i(M) \leq \rho(M) \leq \max_{1\leq i\leq n} r_i(M) = \|M\|_\infty$$
  $$\min_{1\leq j\leq n} c_j(M) \leq \rho(M) \leq \max_{1\leq j\leq n} c_j(M) = \|M\|_1$$

- If $M \geq 0$ does not have a zero row, then $\rho(M) > 0$
- IF $M \geq 0$, $\vec{\mathbf{v}} > 0$, and $M\vec{\mathbf{v}} > c\vec{\mathbf{v}}$ for $c \geq 0$ then $\rho(M) > c$

4.0.2 Perron's Theorem

4.0.2.1 Proposition 1: Let P be a positive matrix.

- If P has a nonnegative eigenvector then this eigenvector and its eigenvalue must in fact be positive
- P can have at most one indepdendent nonnegative eigenvector for any specific eigenvalue.

4.0.2.2 Perron Let $P > 0$. If $P\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ and $\vec{\mathbf{v}} \neq 0$ and $|\lambda| = \rho(P)$ then $\lambda = \rho(P)$, $\vec{\mathbf{v}} = e^{i\theta}|\vec{\mathbf{v}}|$ and $|v| > 0$

4.0.2.3 Perron I Let $P > 0$. Then

- $\rho(P) > 0$
- $\rho(P)$ is an eigenvalue of P
- $\rho(P)$ has a positive eigenvector
- The geometric multiplicity of eigenvalue $\rho(P)$ is 1
- $\rho(P)$ is the only dominant eigenvalue of P

4.0.2.4 Perron Roots: Let $A \in \mathbb{R}^{n\times n}$ be a matrix such that A, A^T both satisfy all conclusions of Perron I. Then $\rho(P)$ is called the **Perron Root** and there exists Perron vectors:

- **(Right) Perron vector:** A unique vector $\vec{\mathbf{z}} > 0$ such that $A\vec{\mathbf{z}} = \rho(A)\vec{\mathbf{z}}$ and $\|\vec{\mathbf{z}}\|_1 = 1$
- **Left Perron vector:** A unique vector $\vec{\mathbf{y}} > 0$ such that $\vec{\mathbf{y}}^T A = \rho(A)\vec{\mathbf{y}}^T$ and $\vec{\mathbf{y}}^T\vec{\mathbf{z}} = 1$

Further, if $\vec{\mathbf{v}} \geq 0$ is an eigenvector of A, then $\vec{\mathbf{v}}$ is a multiple of the Perron vector $\vec{\mathbf{z}}$.

4.0.2.5 Perron II Let $A, A^T \in \mathbb{R}^{n\times n}$ satisfy the conclusions of Perron 1 with Perron vector and $\vec{\mathbf{z}}$ and left Perron vector $\vec{\mathbf{y}}$. (Any positive matrix satisfies these properties.) Then

$$\lim_{k\to\infty}\left(\frac{1}{\rho(A)}A\right)^k = \vec{\mathbf{z}}\vec{\mathbf{y}}^T$$

4.0.2.6 Perron III Let $A \in \mathbb{R}^{n\times n}$ be a matrix such thatthere is a rank one matrix S such that $\lim_{k\to\infty}\left(\frac{1}{\rho(A)}A\right)^k = S$. Then $\rho(A)$ is a simple eigenvalue of A (its algebraic multiplicity is 1)

4.0.3 Regular Matrices

A nonnegative matrix M is **regular** (also known as aperiodic or primitive) if there exists a positive integer k such that $M^k > 0$, the minimum k at which this occurs is called the **exponent**. If $M \geq 0$ is regular of exponent k, then for all $s \geq 0$, $M^{k+s} > 0$.

4.0.3.1 Eventually positive
A real square matrix A is **eventually positive** if there exists $k_0 > 0 \in \mathbb{N}$ such that $k \geq k_0$ implies $A^k > 0$, and k_0 is the **power index** of A.

4.0.4 Perron-Frobenius Property

$A \in \mathbb{R}^{n\times n}$ has the **Strong Perron-Frobenius Property** if the following are satisfied:

- $\rho(A) > 0$ is a simple eigenvalue of A
- $\rho(A)$ is the only dominant eigenvalue of A
- A has a positive eigenvector for $\rho(A)$

For any matrix $A \in \mathbb{C}^{n\times n}$, $\operatorname{spec}(A^k) = \left\{\lambda^k : \lambda \in \operatorname{spec}(A)\right\}$ and if $\vec{\mathbf{v}}$ is an eigenvector of A for eigenvalue λ then $\vec{\mathbf{v}}$ is an eigenvector of A^k for eigenvalue λ^k

4.0.4.1 Theorem Let $A \in \mathbb{R}^{n\times n}$. The following are equivalent:

- A is eventually positive
- A and A^T satisfy the strong Perron-Frodenius property
- A and A^T satisfy the 5 conditions of Perron I.
- $\lim_{k\to\infty}\left(\frac{1}{\rho(A)}A\right)^k = \vec{\mathbf{z}}\vec{\mathbf{y}}^T$ where $Az = \rho(A)\vec{z}$, $\vec{\mathbf{y}}^T A = \rho(A)\vec{\mathbf{y}}^T$, $\vec{\mathbf{z}}, \vec{\mathbf{y}} > 0$ and $\vec{\mathbf{y}}^T\vec{\mathbf{z}} = 1$
- $\lim_{k\to\infty}\left(\frac{1}{\rho(A)}A\right)^k = S > 0$

4.0.5 Digraphs and Reduceable Matrices

A **digraph** $\Gamma = (V, E)$ is a finite nonempty set V of **vertices** and a set of **arcs** $E \subseteq V \times V$.
For $S \subseteq V$ the **subdigraph** of Γ induced by S is digraph $\Gamma[S] := (S, E \cap (S \times S))$.
A digraph of a matrix $A \in \mathbb{C}^{n\times n}$, $\Gamma(A)$ is the graph with vertex set $V(A) = \{1, ..., n\}$ and
LINEAR PROGRESS

5 Functions of Matrices

5.1 Definitions

5.1.1 Jordan Canonical Form definition

For $A \in \mathbb{C}^{n\times n}$ with JCF $Z^{-1}AZ$ where J_A has Jordan blocks $\{J_1(\lambda_1), ..., J_p(\lambda_p)\}$ with $J_k(\lambda_k) \in \mathbb{C}^{m_k\times m_k}$ and $\sum_i m_i = n$ then define

$$f(A) := Zf(J_A)Z^{-1} = Z\mathrm{diag}(f(J_k(\lambda_k)))Z^{-1}$$

where

$$f(J_k(\lambda_k)) := \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \dots & \frac{f^{(m_k-1)(\lambda_k)}}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}$$

- If any of the derivatives is undefined at any of the λ_j, then $f(A)$ is undefined.-is this true of al definitions?
- If f is multivalues and A is derogatory (repeated eigenvaules occurring in more than one Jordan block), the Jordan Canonical Form definition has morethan one interpretaton:
  - **Primary Matrix Function**:= If the same branch for f and its derivatives are taken for each occurence of an eigenvalue in different Jordan blocks. It is usually assumed that $f(A)$ means the primary matrix function.
  - **Nonprimary Matrix Function**:= If different branches for f and its derivatives are taken for the same eigenvalue in two different Jordan blocks. If this is the case, f cannot be expressed as a polynomial of the matrix, and depends on the matrix Z.

5.1.2 Polynomial Interpolation Definition

Given a matrix A and its distinct eigenvalues $\lambda_1, ..., \lambda_s$, let n_i be the **index** of λ_i that is the order of the largest Jordan block in which λ_i appears. Then let $r(A)$ be the unique Hermite interpolating polynomial of degree less than $\sum_{i=n}^{s} n_i$ that satisfies the interpolation conditions $r^{(j)}(\lambda_i) = f^{(j)}(\lambda_j)$ for $j = \{0, 1, ..., n_i - 1\}$ and $i = \{1, ..., s\}$. Then

$$f(A) := r(A)$$

- If any of the derivatives are undefined at any of the λ_j, then $f(A)$ is undefined.

5.1.3 Cauchy Integral Definition

The function f is said to be **defined on the spectrum of** A if all the derivatives exist at all of the eigenvalues λ_j exist. If f is analytic inside a closed contour Γ that encloses $\sigma(A)$, then

$$f(A) := \frac{1}{2\pi i}\int_\Gamma f(z)(zI - A)^{-1}dz$$

Pierson Guthrey
pguthrey@iastate.edu

5.1.4 Matrix Function Facts

- The JCF and Interpolating polynomial definitions are equivalent and the Cauchy Integral definition is equivalent to both if f is anaylitic.
- $f(A)$ is a polynomial in A And the coefficients depend on A
- $f(A)$ commutes with A
- $f(A^T) = f(A)^T$
- For any nonsingular X, $f(XAX^{-1}) = Xf(A)X^{-1}$
- For a block-diagonal matrix $D = \text{diag}(A_1, ..., A_n)$, $f(D) = \text{diag}(f(A_1), ..., f(A_n))$
- If A is diagonalizable with $Z^{-1}AZ = D$ (D diagonal), then $f(A) = Zf(D)Z^{-1}$
- Let f, g be functions defined on the spectrum of A.

$$h(t) = f(t) + g(t) \implies h(A) = f(A) + g(A)$$
$$h(t) = f(t)g(t) \implies h(A) = f(A)g(A)$$

- Let $G(u_1, ..., u_t)$ be polynomial in $u_1, ..., u_n$ and let $f_1, ..., f_n$ be functions defined on the spectrum of A. If $g(\lambda) = G(f_1(\lambda), ..., f_n(\lambda))$ takes zero values on the spectrum of A then $g(A) = G(f_1(A), ..., f_t(A)) = 0$
  Corollary: $\sin^2(A) + \cos^2(A) = I$, $(A^{\frac{1}{p}})^p$, and $e^{iA} = \cos A + i \sin A$
- If f has a Taylor Series expansion with coefficients a_k and radius of convergence r, then for $A \in \mathbb{C}^{n \times n}$

$$f(A) = \sum_{k=0}^{\infty} a_k (A - \alpha I)^k$$

  if and only if each of the distinct eigenvalues $\{\lambda_i\}$ of A satisfy one of the following conditions
  - $|\lambda_i - \alpha| < r$
  - $|\lambda_i - \alpha| = r$ and the series for $f^{n_i - 1}(\lambda)$ is convergent at $\lambda_i \; \forall \; i$
- Let $T \in \mathbb{C}^{n \times n}$ be upper triangular and suppose f is defined on the spectrum of T, then $F = f(T)$ is upper triangular. Let $\lambda_i = t_{ii}$ and let S_{ij} be the set of all strictly increasing sequences of integers that start at i and end at j and let $[\lambda_{s_0}, ..., \lambda_{s_k}]$ be the k^{th} order divided difference of f at $\lambda_{s_0}, ..., \lambda_{s_k}$. Then $f_{ii} = f(t_{ii})$ and

$$f_{ij} = \sum_{(s_0, ..., s_k) \in S_{ij}} t_{s_0, s_1} t_{s_1, s_2} \times ... \times t_{s_0, s_1} t_{s_{k-1}, s_k} f \left[\lambda_{s_0}, ..., \lambda_{s_k}\right]$$

5.2 Specific Functions

5.2.1 Matrix Square Root

Let $A \in \mathbb{C}^{n \times n}$. Any X such that $X^2 = A$ is a **square root** of A

- **Uniqueness**: If A has no eigenvalues on $\mathbb{R}_0^-$ (the closed nagative real axis) then there is a unique square root X of A each of whose eigenvalues are 0 or lies in the right half-plane and $\sqrt{A}$ is a primary matrix function of A. n this case we call X the **principle square root** of A and it is written $X = A^{\frac{1}{2}}$. If A is real then so is $A^{\frac{1}{2}}$. The integral representation of this is

$$f(A) := \frac{2}{\pi} A \int_{\Gamma} (t^2 I + A)^{-1} dt$$

Pierson Guthrey
pguthrey@iastate.edu

- **Existence** A necessary and sufficient condition for a matrix to have a square root is that in the "ascent sequence" of integers $d_1, d_2, ...$ defined by

$$d_i = \dim(\ker(A^i)) - \dim(\ker(A^{i-1}))$$

contains no two identical odd integers.

- $A \in \mathbb{R}^{n\times n}$ has a real square root if and only if A satisfies the "ascent sequence" condition and A has an even number of Jordan blocks of each size for negative eigenvalue.
- A positive semidefinite matrix $A \in \mathbb{C}^{n\times n}$ has a unique positive semidefinite square root
- A singular matrix need not have a square root

5.2.1.1 The Identity matrix

- $I_{n\times n}$ has 2^n diagonal square roots $\mathrm{diag}(\pm 1)$ with primary matrix functions I and $-I$
- Nondiagonal but symmetric nonprimary square roots of I_n including any Householder matrix $I - 2\vec{\mathbf{v}}\vec{\mathbf{v}}^T \setminus (\vec{\mathbf{v}}^T\vec{\mathbf{v}})$ for $\vec{\mathbf{v}} \neq 0$ and the identity with its columns in the reverse order
- Nonsymmetric square roots are easily constructed in the form XDX^{-1} where X is nonsingular but nonorthogonal and $D = \mathrm{diag}(\pm 1) \neq \pm I$

5.2.2 Matrix Exponential

Let $A \in \mathbb{C}^{n\times n}$. Define

$$e^A = exp(A) = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + ... = I + \sum_{k=1}^{\infty} \frac{A^k}{k!}$$

- $e^{(A+B)t} = e^{At}e^{Bt}$ holds for all t if and only if $AB = BA$
- The solution to the differential equaton in $n \times n$ matrices

$$\frac{dY}{dt} = AY + BY, \qquad Y(0) = C, \qquad A, B, Y \in \mathbb{C}^{n\times n}$$

has the solution $Y(t) = e^{At}Ce^{Bt}$

- $A \in \mathbb{C}^{n\times n}$ is unitary if and only if it can be written $A = e^{iH}$ where H is Hermitian. H can be taken to be Hermitian positive definite
- $A \in \mathbb{R}^{n\times n}$ is orthogonal with $\det(A) = 1$ if and only if $A = e^S$ with $S \in \mathbb{R}^{n\times n}$ skew-symmetric.

5.2.3 Matrix Logarithm

Let $A \in \mathbb{C}^{n\times n}$. Any X such that $e^X = A$ is a **logarithm** of A

- **Uniqueness:** If A has no eigenvalues on $\mathbb{R}^-$, then there is a unique logarithm X of A all whose eigenvalues lie in the strip $\{z : -\pi < Im(z) < \pi\}$. This is the **principle logarithm** of A and is written $X = \log A$. If A is real, then $\log(A)$ is real
- **Existence** If $\rho(A) < 1$, then

$$\log(I + A) = A - \frac{A^2}{2} + \frac{A^3}{3} - ... = \sum_{k=1}^{\infty} (-1)^k \frac{A^k}{k}$$

Pierson Guthrey
pguthrey@iastate.edu

- $A \in \mathbb{R}^{n\times n}$ has areal logarithm if and only if A is nonsingular and A has an even number of of Jordan blocks of each size for every negative eigenvalue.
- $e^(\log(A)) = A$ holds when $\log$ is defined on the spectrum of $A \in \mathbb{C}^{n\times n}$, but does not generally hold unless the spectrum of A is restricted.
- If $A \in \mathbb{C}^{n\times n}$ is nonsingular then $\det(A) = e^(\operatorname{tr}(\log(A)))$ where $\log(A)$ is any logarithm of A

For $A \in \mathbb{C}^{n\times m}$

$$\|A\|_1 = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}| \text{ maximum row sum}$$

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{m} |a_{ij}| \text{ maximum column sum}$$

$$\|A\|_2 = \sqrt{\rho(A^*A)}$$

$$\|A\|_F = \sqrt{\operatorname{tr}(A^*A)}$$

For any induced matrix norm $\|\cdot\|$,

$$\rho(A) \le \|A\|$$

If $U \in \mathbb{R}^{n\times n}$ is unitary, then

$$\|Ux\|_2 = \|x\|_2$$

$$\|U\|_2 = 1$$

If $A \in \mathbb{C}^{m\times n}$, then

$$\|AU\|_2 = \|A\|_2$$

also

$$\|AU\|_F = \|A\|_F$$

5.2.3.1 Rayleigh Quotient

$$R_A(\vec{\mathbf{x}}) = \frac{\vec{\mathbf{x}}^* A \vec{\mathbf{x}}}{\vec{\mathbf{x}}^* \vec{\mathbf{x}}}$$

If $A \in \mathbb{C}^{n\times n}$ is Hermitian, then we can compute the greatest eigenvalue

$$\sup_{\vec{\mathbf{x}}\neq 0} R(A) = \max_i \lambda_i$$

For any $A \in \mathbb{C}^{n\times m}$, $\|A\|_2 = \sqrt{\rho(A^*A)}$
For any $A \in \mathbb{C}^{n\times n}$, $\|A\|_2 = \rho(A)$