

Numerical Matrix Analysis

Notes #3 — Orthogonal Vectors, Matrices and Norms

Peter Blomgren

blomgren@sdsu.edu

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

<http://terminus.sdsu.edu/>

Spring 2024

(Revised: January 18, 2024)



Outline

- 1 Student Learning Targets, and Objectives
 - SLOs: Linear Algebra Review, Part II
- 2 Introduction
 - Recap
- 3 Fundamental Concepts
 - Transpose (Adjoint) / Hermitian
 - Inner Products, Matrix Properties, Orthogonality
 - Unitary Matrices, Vector Norms, Matrix Norms
- 4 Next...
 - Looking Ahead



Student Learning Targets, and Objectives

Target Vectors

- Objective Euclidean Inner Product: the Dot Product — Bilinearity
- Objective Orthogonality and Orthonormality
- Objective Linear Independence, Basis
- Objective Projections
- Objective Vector Norms

Target Matrices

- Objective Symmetric and Hermitian Matrices
- Objective Inverses and Hermitian Transposes for Matrix Products
- Objective Unitary Matrices
- Objective Matrix Norms

Target Actions

- Objective Projections



Previously...

A quick review / crash course in basic linear algebra:

- Vectors: Transpose, Addition & Subtraction
- Matrix-Vector Product
- Vandermonde Matrix ... and Linear Least Squares Problems
- Matrix-Matrix Product
- Transpose of a Matrix (A^T)
- Range and Nullspace of a Matrix A
- Rank of a Matrix $A_{m \times n}$
- Inverse of a Matrix A



Now...

...More Fundamental Concepts

The **Transpose** (Adjoint) a.k.a **Hermitian** (Transpose, or Conjugate) of a matrix $A \in \mathbb{C}^{m \times n}$...

For a scalar $z \in \mathbb{C}$, $z = a + bi$, the **complex conjugate** \bar{z} , or z^* is obtained by negating the imaginary part, i.e. $z^* = a - bi$.

Note that if $z \in \mathbb{R}$, then $z^* = z$.

For a matrix $A \in \mathbb{C}^{m \times n}$, the Hermitian Conjugate $A^* \in \mathbb{C}^{n \times m}$ is the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \Rightarrow A^* = \begin{bmatrix} a_{11}^* & a_{21}^* & a_{31}^* & a_{41}^* \\ a_{12}^* & a_{22}^* & a_{32}^* & a_{42}^* \end{bmatrix}$$



The Hermitian Conjugate

If $A = A^*$, the matrix A is said to be **Hermitian**.

Note that a **Hermitian matrix must be square**.

In the case that A is real-valued, i.e. $A \in \mathbb{R}^{m \times n}$, then $A = A^* = A^T$ (the Hermitian conjugate equals the **transpose**).

If $A = A^T$, the matrix A is said to be **Symmetric**.

Our book (TREFETHEN-BAU) tends to state results and theorems in terms of complex vectors and matrices, and hence use the Hermitian conjugate, i.e. \vec{x}^* is a row-vector.

The advantage of this approach is that we are able to state the most general results.

Note: There are some differences in regards to properties over \mathbb{R}^n and \mathbb{C}^n ; those gory details are explored in [MATH 524]



The Inner Product of Two Vectors

a.k.a the **dot product**

The Euclidean **inner product**, denoted $\langle \vec{x}, \vec{y} \rangle$, of two column vectors $\vec{x}, \vec{y} \in \mathbb{C}^m$ is defined

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} = \sum_{i=1}^m x_i^* y_i$$

note that the inner product is a scalar quantity.

The **Euclidean length**, $\|\vec{x}\|$, of $\vec{x} \in \mathbb{C}^m$ is defined

$$\|\vec{x}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{\vec{x}^* \vec{x}} = \sqrt{\sum_{i=1}^m |x_i|^2}$$

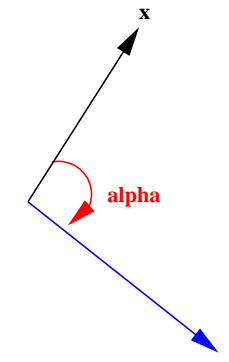


Inner Product: Geometrical Interpretation

The inner product can also be written

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} = \|\vec{x}\| \cdot \|\vec{y}\| \cdot \cos(\alpha)$$

where α is the angle between \vec{x} and \vec{y}



Inner Product: Properties

Bi-Linearity

The inner product is **bilinear**, *i.e.* it is linear in each argument separately:

$$(1) \quad (\vec{x}_1 + \vec{x}_2)^* \vec{y} = \vec{x}_1^* \vec{y} + \vec{x}_2^* \vec{y}$$

$$(2) \quad \vec{x}^* (\vec{y}_1 + \vec{y}_2) = \vec{x}^* \vec{y}_1 + \vec{x}^* \vec{y}_2$$

$$(3) \quad (\alpha \vec{x})^* (\beta \vec{y}) = \alpha^* \beta \vec{x}^* \vec{y}$$

where $\vec{x}, \vec{x}_1, \vec{x}_2, \vec{y}, \vec{y}_1, \vec{y}_2 \in \mathbb{C}^m$, and $\alpha, \beta \in \mathbb{C}$.

Compare: Bilinearity of the matrix-vector product. The Euclidean inner product is really "just" a particular application/interpretation of the matrix-vector product.



Associated Matrix Properties

For any two matrices A and B , of compatible dimensions, *i.e.* $A \in \mathbb{C}^{m \times n}$, and $B \in \mathbb{C}^{n \times k}$ the following holds

$$(AB)^* = B^* A^*$$

If the matrices A and B are square, and invertible, the following holds

$$(AB)^{-1} = B^{-1} A^{-1}$$

When necessary, we use the notation A^{-*} for $(A^*)^{-1} \equiv (A^{-1})^*$.

Question: What is $(AB)^{-*}$ (when well-defined)?



Orthogonal and Orthonormal Vectors

Two vectors are **orthogonal** if and only if $\langle \vec{x}, \vec{y} \rangle = \vec{x}^* \vec{y} = 0$,

$$0 = \frac{\vec{x}^* \vec{y}}{\|\vec{x}\| \cdot \|\vec{y}\|} = \cos(\alpha) \Leftrightarrow \alpha = \pi/2 + k \cdot \pi, \quad k \in \mathbb{Z}.$$

A **set** of **non-zero** vectors S is **orthogonal** if its elements are pairwise orthogonal, *i.e.*

$$\forall \vec{x}, \vec{y} \in S, \quad \vec{x} \neq \vec{y} \Rightarrow \vec{x}^* \vec{y} = 0$$

A **set** of vectors S is **orthonormal** if it is **orthogonal**, and $\forall \vec{x} \in S, \|\vec{x}\| = 1$, *i.e.* all vectors are **unit-vectors**.



Linear Independence of Orthogonal Set

Theorem (Linear Independence)

The vectors in an orthogonal set S are linearly independent.

Proof (Linear Independence of Orthogonal Vectors).

If the vectors in S are not independent, then $\exists \vec{v}_k \in S : \vec{v}_k \neq \vec{0}$, so that

$$\vec{v}_k = \sum_{i \neq k} c_i \vec{v}_i.$$

Since $\vec{v}_k \neq \vec{0}$, $\langle \vec{v}_k, \vec{v}_k \rangle > 0$, now we use the bi-linearity property of inner products, and the orthogonality of S :

$$0 < \langle \vec{v}_k, \vec{v}_k \rangle = \left\langle \vec{v}_k, \sum_{i \neq k} c_i \vec{v}_i \right\rangle = \sum_{i \neq k} c_i \underbrace{\langle \vec{v}_k, \vec{v}_i \rangle}_{0, \forall i \neq k} = 0.$$

This contradicts the assumption that the vectors are linearly dependent, hence proving the theorem. \square



Corollary: Basis for \mathbb{C}^m

Corollary

If an orthogonal set $S \subseteq \mathbb{C}^m$ contains m vectors, then it is a basis for \mathbb{C}^m .

We can write any vector $\vec{v} \in \mathbb{C}^m$ as a unique linear combination

$$\vec{v} = \sum_{i=1}^m a_i \vec{s}_i, \quad \text{where } a_i = \frac{\langle \vec{s}_i, \vec{v} \rangle}{\|\vec{s}_i\|^2}.$$

The computation of $a_i \vec{s}_i$ is a **projection** of \vec{v} onto the direction \vec{s}_i .

We can use this in order to decompose arbitrary vectors into orthogonal components...



Orthogonal Vector Components

1 of 3

Suppose we have an **orthonormal set** of vectors $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$, $\vec{q}_i \in \mathbb{C}^m$, $n \leq m$.

Now, for any vector $\vec{v} \in \mathbb{C}^m$, the vector

$$\vec{r} = \vec{v} - \sum_{i=1}^n \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i$$

is orthogonal to $\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n\}$:

$$\langle \vec{q}_k, \vec{r} \rangle = \langle \vec{q}_k, \vec{v} \rangle - \underbrace{\sum_{i=1}^n \langle \vec{q}_i, \vec{v} \rangle \langle \vec{q}_k, \vec{q}_i \rangle}_{\langle \vec{q}_k, \vec{v} \rangle \underbrace{\langle \vec{q}_k, \vec{q}_k \rangle}_1} = 0.$$



Orthogonal Vector Components

2 of 3

We see that by applying this procedure, we have decomposed the vector \vec{v} into $(n + 1)$ orthogonal components:

$$\vec{v} = \vec{r} + \sum_{i=1}^n \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i$$

If $\{\vec{q}_i\}$ is a basis for \mathbb{C}^m , then $n = m$ and $\vec{r} = \vec{0}$, i.e.

$$\vec{v} \stackrel{!}{=} \sum_{i=1}^n \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i = \sum_{i=1}^n (\vec{q}_i^* \vec{v}) \vec{q}_i = \sum_{i=1}^n \vec{q}_i (\vec{q}_i^* \vec{v}) = \sum_{i=1}^n (\vec{q}_i \vec{q}_i^*) \vec{v}$$



Orthogonal Vector Components

3 of 3

$$\vec{v} = \sum_{i=1}^n \langle \vec{q}_i, \vec{v} \rangle \vec{q}_i = \sum_{i=1}^n (\vec{q}_i^* \vec{v}) \vec{q}_i = \sum_{i=1}^n \vec{q}_i (\vec{q}_i^* \vec{v}) = \sum_{i=1}^n (\vec{q}_i \vec{q}_i^*) \vec{v}$$

In the expression $\sum_{i=1}^n (\vec{q}_i^* \vec{v}) \vec{q}_i$ we view \vec{v} as a linear combination of the vectors \vec{q}_i , with coefficients $(\vec{q}_i^* \vec{v})$; whereas in the mathematically equivalent expression $\sum_{i=1}^n (\vec{q}_i \vec{q}_i^*) \vec{v}$, we view \vec{v} as a sum of **orthogonal projections** onto the various directions \vec{q}_i .

We will return to the issue of projection matrices of the form formed by other products, $\vec{q}_i \vec{q}_i^*$, soon.



Unitary Matrices

A square matrix $Q \in \mathbb{C}^{m \times m}$ is **unitary** (in the real case “orthogonal”) if

$$Q^* = Q^{-1} \Leftrightarrow Q^* Q = I$$

In terms of the columns, \vec{q}_i of Q this looks like

$$\begin{bmatrix} \text{---} & \vec{q}_1^* & \text{---} \\ \text{---} & \vec{q}_2^* & \text{---} \\ & \vdots & \\ \text{---} & \vec{q}_m^* & \text{---} \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

We have $\vec{q}_i^* \vec{q}_j = \delta_{ij}$, the **Kronecker delta**, equal to 1 if and only if $i = j$, and 0 otherwise.



Multiplication by a Unitary Matrix

Since the norms of the columns of a unitary matrix are 1, multiplication by a unitary matrix preserves the Euclidean norm, and inner product in the following sense:

For a unitary Q :

$$(1) \quad \langle Q\vec{x}, Q\vec{y} \rangle = (Q\vec{x})^* (Q\vec{y}) = \vec{x}^* \underbrace{Q^* Q}_I \vec{y} = \vec{x}^* \vec{y} = \langle \vec{x}, \vec{y} \rangle$$

$$(2) \quad \|Q\vec{x}\| = \|\vec{x}\|$$

The invariance of inner products mean that angles between vectors are preserved.

In the real case, multiplication by an orthogonal matrix corresponds to a **rigid rotation** (if $\det(Q) = 1$) or a combined **rotation – reflection** (if $\det(Q) = -1$) of the vector space.



Vector Norms

Norms give us the essential notion of size and distance in a vector space — these are our tools for measuring the quality of approximations and convergence in our algorithms.

Definition (Norm)

A **norm** is a function $\|\cdot\| : \mathbb{C}^m \rightarrow \mathbb{R}$ that assigns a real-valued (length) to each vector. A norm must satisfy the following three conditions for all vectors $\vec{x}, \vec{y} \in \mathbb{C}^m$, and scalars $\alpha \in \mathbb{C}$,

- (1) $\|\vec{x}\| \geq 0$, and $\|\vec{x}\| = 0$ only if $\vec{x} = 0$
- (2) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$
- (3) $\|\alpha\vec{x}\| = |\alpha| \|\vec{x}\|$

(2) is known as the “**triangle inequality**.”



The p -norms

1 of 3

The p -norms (sometimes referred to as the ℓ_p -norms), parametrized by p are defined by

$$\|\vec{x}\|_p = \left[\sum_{i=1}^m |x_i|^p \right]^{1/p}$$

As an illustration, the **unit sphere** $\|\vec{x}\|_p = 1$, $\vec{x} \in \mathbb{R}^2$ is illustrated for some common (and uncommon) p -norms, on the following slides.

2-norm the standard Euclidean length function.

1-norm sometimes referred to as the Manhattan/taxicab-distance.

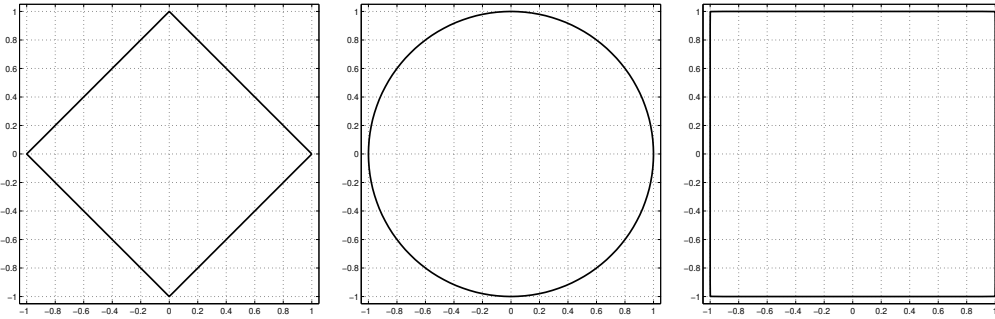
0-norm counts the number of non-zero elements in a vector.



The p -norms

2 of 3

Some commonly used p -norms



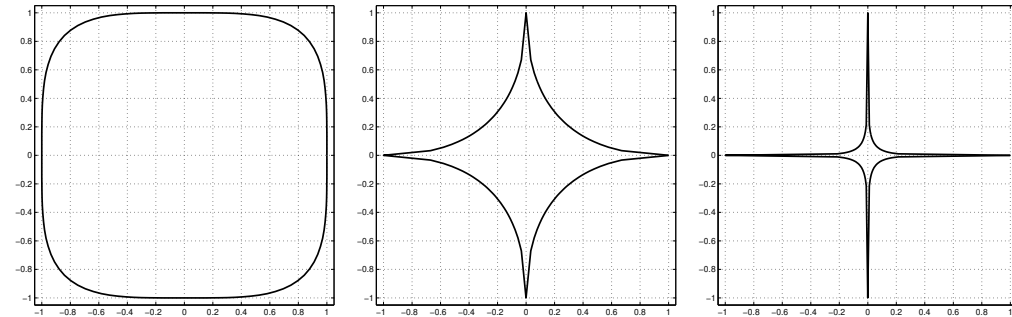
$$\|\vec{x}\|_1 = \sum_{i=1}^m |x_i|, \quad \|\vec{x}\|_2 = \left[\sum_{i=1}^m |x_i|^2 \right]^{1/2}, \quad \|\vec{x}\|_\infty = \max_{i=1 \dots m} |x_i|$$



The p -norms

3 of 3

Some exotic p -{norms,non-norms}



$$\|\vec{x}\|_4 = \left[\sum_{i=1}^m |x_i|^4 \right]^{1/4}, \quad \|\vec{x}\|_{1/2} = \left[\sum_{i=1}^m |x_i|^{1/2} \right]^2, \quad \|\vec{x}\|_{1/4} = \left[\sum_{i=1}^m |x_i|^{1/4} \right]^4$$

Note: when $p < 1$ the “norms” are not convex; which means the triangle inequality will not hold; and strictly speaking these are not norms...

☞ Movie.



Weighted p -norms

1 of 3

The **weighted p -norms** $\|\cdot\|_{W,p}$ are derived from the p -norms:

$$\|\vec{x}\|_{W,p} = \|W\vec{x}\|_p$$

where W is e.g. a diagonal matrix, in which the i th diagonal entry is the weight $w_i \neq 0$:

$$\|\vec{x}\|_{W,p} = \left[\sum_{i=1}^m |w_i x_i|^p \right]^{1/p}$$



Weighted p -norms

2 of 3

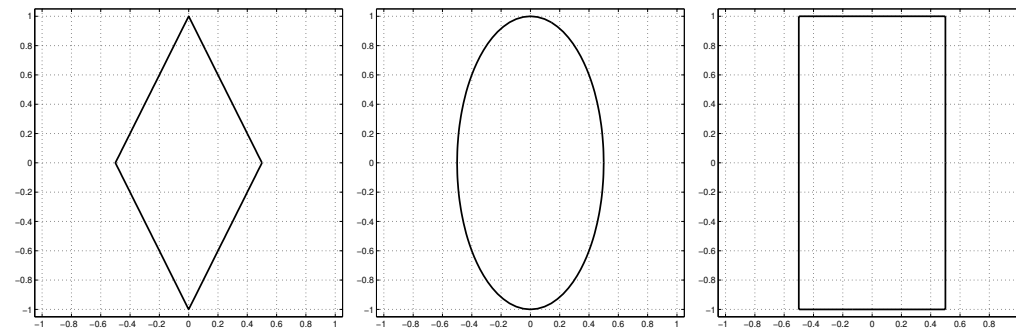


Figure: Visualization of the unit-sphere for the weighted 1-, 2- and ∞ -norms, where $W = \text{diag}(2, 1)$.

The concept of weighted p -norms can be generalized to arbitrary non-singular weight matrices W .



Weighted p -norms

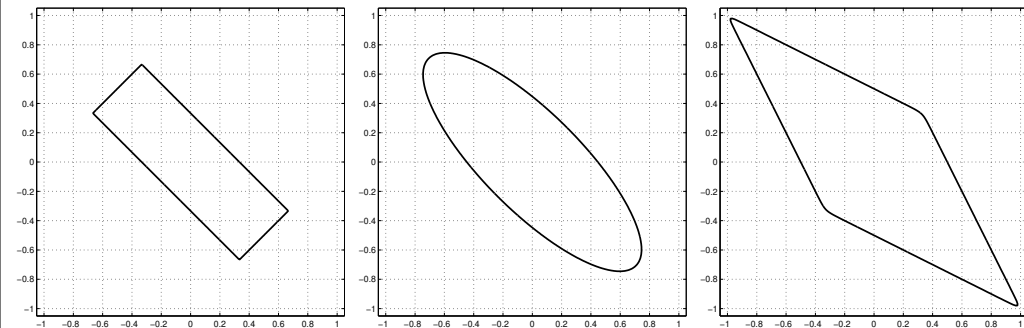


Figure: Visualization of the unit-sphere for the weighted 1-, 2- and ∞ -norms, where $W = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

∃ Movie.



Matrix Norms — Induced by Vector Norms

Given a vector norms $\|\cdot\|_{(m)}$ and $\|\cdot\|_{(n)}$ on the domain and range of $A \in \mathbb{C}^{m \times n}$, the induced matrix norm $\|A\|_{(m,n)}$ is

$$\|A\|_{(m,n)} = \sup_{\vec{x} \in \mathbb{C}^n - \{\vec{0}\}} \left[\frac{\|A\vec{x}\|_{(m)}}{\|\vec{x}\|_{(n)}} \right]$$

In any sane application, both $\|\cdot\|_{(m)}$ and $\|\cdot\|_{(n)}$ will be of the same type, *i.e.* the p -norms (with the same p).

Due to the linearity of norms — the third norm-condition — it is sufficient to maximize the matrix norm over $\vec{x} \in \mathbb{C}^n : \|\vec{x}\| = 1$...

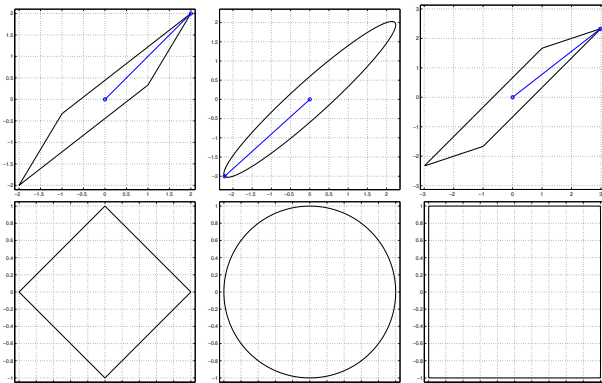
Most of the time the norms with $p = 2$ are used. Indeed, if nothing else is specified, this is usually implied.



Illustration: Matrix Norms

$$A = \begin{bmatrix} 1 & 2 \\ 1/3 & 2 \end{bmatrix}, \quad \lambda(A) = \{2.45743, 0.54257\} \text{ eigenvalues}$$

$$\sigma(A) = \{2.98523, 0.44664\} \text{ singular values}$$



$$\|A\|_1 = 4 \quad \|A\|_2 \approx 2.9852 \quad \|A\|_\infty = 3$$



Special Cases: Matrix p -norms

If D is a diagonal matrix, then

$$\|D\|_p = \max_{1 \leq i \leq m} |d_i|.$$

The 1-norm of a matrix is the maximal column-abs-sum:

$$\|A\|_1 = \max_{1 \leq j \leq n} \|\vec{a}_j\|_1$$

The ∞ -norm of a matrix is the maximal row-abs-sum:

$$\|A\|_\infty = \max_{1 \leq i \leq m} \|\vec{a}_i^*\|_1$$



Next Time

- Additional discussion on norms:
 - Inequalities, General matrix norms, The Frobenius norm, Bounds on norms of products of matrices.
- **The Singular Value Decomposition (SVD).**

