Set 1: Basics of Vector Spaces

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 1
Fall 2017

Scalars, Vectors and Matrices

Scalars and their operations are assumed to be from

- the field of real numbers ($\mathbb{R}$)
- the field of complex numbers ($\mathbb{C}$)
  - complex number: $\alpha = \beta + i\gamma$ where i here is used to represent the root of -1 (occasionally we will use j for this but it will be made clear when this is done)
  - β and γ are the real and imaginary parts of α respectively
  - complex conjugate $\bar{\alpha} = \beta - i\gamma$
  - the absolute value of α denoted $|\alpha|$ is $\sqrt{\alpha\bar{\alpha}} = \sqrt{\beta^2 + \gamma^2}$

Scalars, Vectors and Matrices

- $\mathbb{R}^n$ – a vector is an one-dimensionally ordered list of n real scalars
  - addition of vectors is componentwise scalar addition
  - scalar vector product multiplies each component of the vector with the scalar
- $\mathbb{C}^n$ – a vector is an one-dimensionally ordered list of n complex scalars
  - addition of vectors is componentwise complex scalar addition
  - scalar vector product multiplies each complex component of the vector with the complex scalar

Example – $\mathbb{R}^3$

Vectors:

$$x = \begin{pmatrix} 1 \\ 3 \\ -52 \end{pmatrix} \quad y = \begin{pmatrix} 10 \\ -4 \\ 2 \end{pmatrix}$$

Basic Operations:

$$x + y = \begin{pmatrix} 11 \\ -1 \\ -50 \end{pmatrix} \quad 2x = \begin{pmatrix} 2 \\ 6 \\ -104 \end{pmatrix} \quad 3y = \begin{pmatrix} 30 \\ -12 \\ 6 \end{pmatrix}$$

Linear Combination:

$$2x + 3y = \begin{pmatrix} 32 \\ -6 \\ -98 \end{pmatrix}$$

Example – $\mathbb{R}^3$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad e = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Scalars, Vectors and Matrices

Definition 1.1. An $m \times n$ matrix of scalars from $\mathbb{R}$ or $\mathbb{C}$ is a two-dimensionally ordered arrangement of mn scalars

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

The set of $m \times n$ matrices with scalar elements from $\mathbb{R}$ is denoted $\mathbb{R}^{m \times n}$

The set of $m \times n$ matrices with scalar elements from $\mathbb{C}$ is denoted $\mathbb{C}^{m \times n}$

Matrix Operations

Matrix scaling $A, B \in \mathbb{R}^{m \times n}$ and $\gamma \in \mathbb{R}$:

$$B = \gamma A = A\gamma \text{ has elements } \beta_{ij} = \gamma\alpha_{ij}$$

Matrix addition $A, B, C \in \mathbb{R}^{m \times n}$:

$$C = A + B = B + A \text{ has elements } \gamma_{ij} = \beta_{ij} + \alpha_{ij}$$

This is the collection of vectors $\mathbb{R}^{mn}$ and the associated scalar field and operations

Matrix Vector Product

Definition 1.2. If

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \in \mathbb{R}^{m \times n}$$

and the vector $x \in \mathbb{R}^n$

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

then

$$Ax = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$$

Matrix Multiplication

If $A \in \mathbb{R}^{n_1 \times n_2}$, $B \in \mathbb{R}^{n_2 \times n_3}$, then $C \in \mathbb{R}^{n_1 \times n_3}$ is

Scalar definintion:

$$C = AB \text{ has elements } \gamma_{ij} = \Sigma_{k=1}^{n_2} \alpha_{ik} \beta_{kj}$$

Matrix-vector definition:

$$C = AB \rightarrow c_i = Ab_i \quad i = 1, \ldots, n_3 \text{ where } c_i = Ce_i, \quad b_i = Be_i$$

Outer product definition:

$$C = AB = \Sigma_{i=1}^{n_2} a_i b_i^T \text{ where } a_i = Ae_i, \quad b_i^T = e_i^T B$$

Inner product definintion:

$$C = AB \text{ has elements } \gamma_{ij} = a_i^T b_j \text{ where } b_i = Be_i, \quad a_i^T = e_i^T A$$

Matrix Multiplication

- matrix multiplication definitions are consistent with matrix-vector product definitions
- the matrix product is not commutative
- the matrix product is associative
- the matrix product is distributive, i.e., $A(B+C) = AB + AC$
- All scalars and vectors can be replaced with submatrices of appropriate dimension to yield block forms of the matrix product

Matrix Operations

Definition 1.3. The transpose of $A \in \mathbb{R}^{m\times n}$, denoted A^T, and the hermitian transpose of $A \in \mathbb{C}^{m\times n}$, denoted A^H, are the $n \times m$ matrices

$$A^T = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{m2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{mn} \end{pmatrix} \quad A^H = \begin{pmatrix} \bar{\alpha}_{11} & \bar{\alpha}_{21} & \cdots & \bar{\alpha}_{m1} \\ \bar{\alpha}_{12} & \bar{\alpha}_{22} & \cdots & \bar{\alpha}_{m2} \\ \vdots & \vdots & & \vdots \\ \bar{\alpha}_{1n} & \bar{\alpha}_{2n} & \cdots & \bar{\alpha}_{mn} \end{pmatrix}$$

Vector Space

Definition 1.4. Given scalars $\mathcal{F}$, a set of vectors $\mathcal{V}$, a vector addition operation $x = y + z$ for $x, y, z \in \mathcal{V}$, and a scalar-vector product operation $y = \alpha x$ for $x, y \in \mathcal{V}$ and $\alpha \in \mathcal{F}$, we have a vector space if the following properties hold:

$$x + y = y + x \tag{1}$$

$$(x + y) + z = x + (y + z) \tag{2}$$

$$x + 0_v = x \tag{3}$$

$$x + (-1_s)x = 0_v \tag{4}$$

$$(\alpha\beta)x = \alpha(\beta x) \tag{5}$$

$$(\alpha +_s \beta)x = \alpha x + \beta x \tag{6}$$

$$\alpha(x + y) = \alpha x + \alpha y \tag{7}$$

$$1_s x = x \tag{8}$$

Scalar and Vector 0

$$
\begin{aligned}
0_v &= a + (-1)a \quad prop4 \\
&= 1a + (-1)a \quad prop8 \\
&= (0+1)a + (-1)a \quad scalar\ 0+1=1 \\
&= (0a + 1a) + (-1)a \quad prop6 \\
&= 0a + (1a + (-1)a) \quad prop2 \\
&= 0a + (a + (-1)a) \quad prop8 \\
&= 0a + (0) \quad prop4 \\
&= 0a \quad prop3
\end{aligned}
$$

Examples

- $\mathcal{P}_n$ – the set of polynomials of degree less than or equal to n
  - isomorphic to $\mathbb{C}^{n+1}$
  - elements can be written as a linear combination of $n+1$ monomials therefore finite dimensional space
- $\mathcal{P}_\infty$ – the set of polynomials of any degree
  - any element can be written as a finte sum of monomials
  - infinite dimensional since it is not the same finite sum size for all vectors
- $\mathcal{L}^2_\omega[\alpha, \beta] = \{f : [\alpha, \beta] \to \mathbb{R}, \;\; \int_\alpha^\beta f^2(x)\omega(x)dx < \infty\}$
  - infinite dimensional
  - need concept of convergence to discuss infinite linear combination that represents each vector

Algebraic Structure

- The algebraic structure of a vector space considers:
  - Subspaces
  - Linear Transformations
  - Bases
  - Linear Independence
- The algebraic structure of the vector spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ is **common to all finite dimensional vector spaces**. We will use $\mathbb{R}^n$ in most of our discussions but the results can be adapted to $\mathbb{C}^n$ and all other such vector spaces.
- By definition a vector space $\mathcal{V}$ is closed under linear combinations, but an arbitrary subset of the space is not necessarily closed, e.g., a finite set or the set of vectors with nonnegative elements.

Subspace

Definition 1.5. A subset $\mathcal{S} \subseteq \mathbb{R}^n$ is a **subspace** if it is closed under linear combination, i.e., if $x_1, x_2, \ldots, x_k \in \mathcal{S}$ then for any scalars α_i, $i = 1, \ldots, k$

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \in \mathcal{S}$$

and in fact the subspace is itself a vector space (and hence all of our results apply within $\mathcal{S}$).

Definition 1.6. Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a subset (finite or infinite). The set of all linear combinations of vectors in $\mathcal{S}$ is called the **span** of $\mathcal{S}$ and is a subspace.

Example 1.1. $\mathbb{R}^n = span(e_1, e_2, \ldots, e_n)$

Matrices and Transformations

Definition 1.7. Given $A \in \mathbb{C}^{m \times n}$, consider $b = Ax$ for all $x \in \mathbb{C}^n$.

- The span of the columns of A is a subspace of $\mathbb{C}^m$ called the **range** of A and is denoted $\mathcal{R}(A)$.
- Since $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$, A defines a linear function

$$F_A : \mathbb{C}^n \to \mathcal{R}(A) \subseteq \mathbb{C}^m : x \mapsto Ax$$

- Any linear function $F : \mathbb{C}^n \to \mathbb{C}^m$ has a unique A defining it.
- Matrix multiplication $\Leftrightarrow$ linear function composition

$$F_B \circ F_A \Leftrightarrow C = BA$$

$$y = F_B(F_A(x)) \Leftrightarrow y = B(Ax) = (BA)x = Cx$$

Independence

Definition 1.8. The set of vectors $x_1, \ldots, x_k$ are **linearly independent** if

$$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0 \rightarrow \alpha_i = 0$$

for $i = 1, \ldots, k$. If this does not hold then the vectors are **linearly dependent**.

Note that:

- A set of vectors being linearly dependent implies one of the vectors can be written as a linear combination of the others.
- Any set that contains the 0 vector is linearly dependent.

Examples

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

are linearly independent in $\mathbb{R}^3$.

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad z = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$$

are linearly dependent

Bases

Definition 1.9. A set of vectors $x_1, x_2, \ldots, x_k \in \mathcal{S} \subseteq \mathbb{R}^n$ is a **basis** for the subspace $\mathcal{S}$ if

- $x_1, x_2, \ldots, x_k$ are linearly independent,
- $span(x_1, x_2, \ldots, x_k) = \mathcal{S}$

Note that:

- A subspace has many bases but every basis contains k vectors and the unique integer k is the dimension of the subspace ($k = dim(\mathcal{S})$).
- $k = dim(\mathcal{S})$ is the number of degrees of freedom in $\mathcal{S}$, i.e., $\mathcal{S}$ is essentially $\mathbb{R}^k$ embedded in $\mathbb{R}^n$.
- Any collection of vectors in $\mathcal{S}$ with $k+1$ or more vectors is linearly dependent.

Matrix Implications

- Linear independent columns of $A \in \mathbb{C}^{m\times n} \leftrightarrow \forall x \neq 0,\ Ax \neq 0$
- Linear dependent columns of $A \in \mathbb{C}^{m\times n} \leftrightarrow \exists x \neq 0\ \ni Ax = 0$
- $\mathcal{N}(A) = \{x \in \mathcal{C}^n | Ax = 0\}$ is a subspace called the **null space** of A. (Also called the kernel denoted $\ker(A)$.)
- $\#$ of independent columns =dimension of $\mathcal{R}(A)$ = **column rank** of A
- $\#$ of independent rows =dimension of $\mathcal{R}(A)$ = **row rank** of A
- If $b = Ax \in \mathcal{R}(A)$ and $rank(A) = n$ then the linear function defined by A is one-to-one and onto $\mathcal{R}(A)$ and x is unique.

Analytic Properties

In addition to the algebraic properties discussed so far we can also define analytic properties of vector spaces and the associated linear transformations,

- size
- distance
- angle

These are analyzed via:

- norms
- inner products

Size and Distance

Definition 1.10. A vector norm, $\|x\|$, is a function $\mathbb{C}^n \rightarrow \mathbb{R}$ that satisfies

- $\|x\| \geq 0$ and $x = 0 \leftrightarrow \|x\| = 0$ (definiteness)
- $\|\alpha x\| = |\alpha| \|x\|$ (homogeneity)
- $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)

We can also deduce

$$\|x - y\| \geq |\|x\| - \|y\||$$

Examples Vector Norms

Let $x \in \mathbb{C}^n$ with elements $e_i^H x = \xi_i$.

$$
\begin{aligned}
\|x\|_1 &= \Sigma_{i=1}^n |\xi_i| \\
\|x\|_2 &= \sqrt{\Sigma_{i=1}^n |\xi_i|^2} \\
\|x\|_p &= (\Sigma_{i=1}^n |\xi_i|^p)^{1/p} \\
\|x\|_\infty &= \max_{1 \le i \le n} |\xi_i|
\end{aligned}
$$

Norm Equivalence

Theorem 1.1. *Let $\mu(x)$ and $\nu(x)$ be vector norms then there exist constants, i.e., independent of x, $\sigma > 0$ and $\tau > 0$ such that*

$$\sigma\mu(x) \leq \nu(x) \leq \tau\mu(x)$$

Norm Equivalence

In other words, for analytical purposes, all norms are equivalent. Convergence in one vector norm implies convergence in any other.

Note that σ and τ may be dependent on n.

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

Matrix Norms

Definition 1.11. A matrix norm on $\mathbb{C}^{m\times n}$ denoted $\|A\|$ maps $\mathbb{C}^{m\times n} \to \mathbb{R}$ and satisfies

- $\|A\| \geq 0$ and $A = 0 \leftrightarrow \|A\| = 0$
- $\|\alpha A\| = |\alpha|\|A\|$
- $\|A + B\| \leq \|A\| + \|B\|$

Examples of matrix norms

Let $A \in \mathbb{C}^{m \times n}$ with elements $e_i^H A e_j = \alpha_{ij}$.

$$
\begin{aligned}
\|A\|_1 &= \max_{1 \le j \le n} \Sigma_{i=1}^{m} |\alpha_{ij}| = \max_{1 \le j \le n} \|Ae_j\|_1 \\
\|A\|_\infty &= \max_{1 \le i \le m} \Sigma_{j=1}^{n} |\alpha_{ij}| = \max_{1 \le i \le m} \|e_i^H A\|_1 \\
\|A\|_2 &= \max_{\|x\|_2 = 1} \|Ax\|_2 \\
\|A\|_p &= \max_{\|x\|_p = 1} \|Ax\|_p \\
\|A\|_F &= \sqrt{\Sigma_{i=1}^{m} \Sigma_{j=1}^{n} |\alpha_{ij}|^2} = \sqrt{\Sigma_{i=1}^{n} \|Ae_i\|_2^2}
\end{aligned}
$$

Examples of matrix norms

- The Frobenius norm $\|A\|_F$ is the vector 2-norm applied to the matrix as if it was a element of $\mathbb{C}^{mn}$. Such a matrix norm can be defined based on any vector p-norm.
- $\|A\|_F^2 = trace(A^H A)$ where the trace is the sum of the diagonal elements.
- The definition of $\|A\|_2$ the given requires optimization

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$$

  but $\|A\|_2$ also can be related to eigenvalues and singular values but these are also "infinite" computations.
- $\|A\|_2$ is also called the spectral norm.

Examples of matrix norms

- The Schatten p-norms are a family of matrix norms that are defined by applying a vector p-norm to the vector containing the singular values of the matrix, i.e.,

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

- These are not the same as the induced matrix p-norms.
- $p = 1$ is called the nuclear norm, $\|A\|_*$.
- $p = 2$ is the same as the Frobenius norm.

Equivalence and Differences

- Bounds can be derived in terms of $\|A\|_1$ and $\|A\|_\infty$, i.e., equivalence can be used for approximation

$$\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$$

- While all matrix norms are equivalent for analytical purposes, they **differ considerably in their ease of computation**.

Consistent Matrix Norms

Definition 1.12. The matrix norms $\|\cdot\|_\alpha$,$\|\cdot\|_\beta$,$\|\cdot\|_\gamma$ are **consistent** if

$$\|AB\|_\alpha \leq \|A\|_\beta \|B\|_\gamma$$

whenever the product exists. Note the order of the product and the association of the norms with each term, i.e., β with A, γ with B, matters in this result.

Lemma 1.2. *The matrix p-norm defines a family of consistent matrix norms. Specifically, for $A \in \mathbb{C}^{m\times n}$,$B \in \mathbb{C}^{n\times r}$ and $x \in \mathbb{C}^n$*

$$\|AB\|_p \leq \|A\|_p \|B\|_p$$

$$\|Ax\|_p \leq \|A\|_p \|x\|_p$$

Induced Matrix Norms

Definition 1.13. The matrix norm $\|\cdot\|$ is **subordinate** to vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ if

$$\|Ax\|_\alpha \leq \|A\|\|x\|_\beta$$

and the matrix norm therefore bounds the expansion/contraction of the linear transformation defined by A.

This motivates the following definition.

Definition 1.14. Given vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ the induced matrix norm $\|\cdot\|_{\alpha,\beta}$ is

$$\|A\|_{\alpha,\beta} = \max_{\|x\|_\alpha=1} \|Ax\|_\beta$$

Induced Matrix Norms

Theorem 1.3. *Given a vector norm* $\|\cdot\|_\alpha$ *on* $\mathbb{C}^n$ *or* $\mathbb{R}^n$ *the induced matrix norm* $\|\cdot\|_\beta$ *for an* $n \times n$ *matrix*

1. $\|Ax\|_\alpha \leq \|A\|_\beta \|x\|_\alpha$ *(subordinate)*
2. $\|I\|_\beta = 1$
3. $\|AB\|_\beta \leq \|A\|_\beta \|B\|_\beta$ *(submultiplicative)*

- Note that some texts use the term subordinate norm rather than induced norm.
- Some texts require a matrix norm for square matrices to be submultiplicative, e.g., Horn and Johnson's Matrix Theory.

Convergence

Both vector sequences and matrix sequences can therefore be said to converge to limit vectors and limit matrices by considering convergence in $\mathbb{R}$.

Definition 1.15. For the vector sequence $\{x_k\}$ and the matrix sequence $\{A_k\}$

$$\lim_{k\to\infty} x_k = x \leftrightarrow \lim_{k\to\infty} \|x_k - x\| = 0$$

$$\lim_{k\to\infty} A_k = A \leftrightarrow \lim_{k\to\infty} \|A_k - A\| = 0$$

Componentwise convergence for both follows.

Angles in n-dimensional Spaces

Definition 1.16. An inner product (or scalar product) on a vector space $\mathcal{V}$ is a map $< \cdot, \cdot >: \mathcal{V} \times \mathcal{V} \to F$ where the field F is either $\mathbb{R}$ or $\mathbb{C}$ that satisfies

1. $< \alpha x + \beta z, y > = \alpha < x, y > + \beta < z, y >$, with $x, y, z \in \mathcal{V}$ and $\alpha, \beta \in F$. (linearity)

2. $< x, y > = \overline{< y, x >}$ (hermitian)

3. $< x, x > \geq 0$ and $< x, x > = 0 \leftrightarrow x = 0$ (definiteness)

Inner Product

- $< x, y >= x^H y$ is an inner product for $\mathbb{C}^n$
- $< x, y >= x^T y$ is an inner product for $\mathbb{R}^n$
- There are other inner products for $\mathbb{C}^n$ and $\mathbb{R}^n$.
- $\|x\| = \sqrt{< x, x >}$ is a norm.

Angles in n-dimensional Spaces

Lemma 1.4. *For* $x, y \in \mathbb{C}^n$

- $|x^H y| \leq \|x\|_p \|y\|_q$ *with* $\frac{1}{p} + \frac{1}{q} = 1$ *(Hoelder inequality)*
- $|x^H y| \leq \|x\|_2 \|y\|_2$ *(Cauchy-Schwarz inequality)*
- $|x^H y| \leq \|x\|_1 \|y\|_\infty$

Angles in n-dimensional Spaces

Angles can be defined by making the Cauchy-Schwarz inequality an equality.

Definition 1.17. Let x and y be two nonzero vectors in $\mathbb{C}^n$ then the cosine of the **angle between the one-dimensional subspaces defined by the vectors**, $0 \leq \theta \leq \pi/2$, is defined

$$|x^H y| = cos\theta \|x\|_2 \|y\|_2$$

Definition 1.18. Let x and y be two nonzero vectors in $\mathbb{C}^n$ then the cosine of the **angle between the vectors**, $0 \leq \theta < 2\pi$ or $-\pi \leq \phi \leq \pi$, is defined

$$x^H y = cos\theta \|x\|_2 \|y\|_2 = cos\phi \|x\|_2 \|y\|_2$$

Generalization from $\mathbb{R}^2$

Consider $x, y \in \mathbb{R}^2$ positive quadrant.

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

$$\tilde{x}^T \tilde{y} = \cos\theta$$

$$\tilde{x} = (\cos\theta_1, \sin\theta_1) \text{ and } \|\tilde{x}\| = 1$$

$$\tilde{y} = (\cos\theta_2, \sin\theta_2) \text{ and } \|\tilde{y}\| = 1$$

where θ_1 and θ_2 are angles from $(1,0)$

$$\tilde{x}^T \tilde{y} = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 = \cos(\theta_1 - \theta_2) = \cos\theta$$

Orthogonality

Definition 1.19. The vectors x and y are said to be orthogonal if their inner product is 0, i.e., $< x, y >= x^H y = 0$.

This generalizes the Pythagorean Theorem to multidimensional and complex vectors:

$$\begin{aligned} \|x + y\|_2^2 &= (x+y)^H (x+y) \\ &= x^H x + y^H y + 2Re(x^H y) \\ &= x^H x + y^H y \\ &= \|x\|_2^2 + \|y\|_2^2 \end{aligned}$$

Polarization and Parallelograms

Theorem 1.5. *Let $\mathcal{V}$ be a vector space over $\mathbb{R}$ (similar statments can be made for $\mathbb{C}$) with and inner product $< x, y >$. If the norm is defined by$\|x\| = \sqrt{< x, x >}$ then we have*

- $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 < x, y >$
- $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2 < x, y >$ *(cosine law)*
- $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ *(parallelogram law)*
- $\|x + y\|^2 - \|x - y\|^2 = 4 < x, y >$ *(polarization identity)*

Polarization and Parallelograms

The reverse is also true.

Theorem 1.6. *Let $\mathcal{V}$ be a normed vector space over $\mathbb{R}$ (similar statments can be made for $\mathbb{C}$). If the norm satisfies the parallelogram law then the polarization identity defines an inner product for $\mathcal{V}$. That is*

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

$$\Downarrow$$

$$< x, y >= \frac{1}{4}\{\|x+y\|^2 - \|x-y\|^2\}$$

$$\Downarrow$$

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2 < x, y >$$

$$\textit{and } \|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2 < x, y >$$