The Vector Norm

Definition Let V be a vector space. A *norm* on V is a real-valued function $||\cdot||: V \to \mathbb{R}$ that satisfies the following properties.

1. $||\mathbf{v}|| \geq 0$ for all $\mathbf{v} \in V$ and $||\mathbf{v}|| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

2. $||\alpha\ \mathbf{v}|| = |\alpha|\ ||\mathbf{v}||$ for all scalars α and all vectors $\mathbf{v} \in V$.

3. $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$ for all $\mathbf{u}$, $\mathbf{v} \in V$.

A vector $\mathbf{v}$ is said to be a *normal vector* if $||\mathbf{v}|| = 1$.

The Inner Product

Definition Let V be a real vector space. A (real) inner product on V is a function $(\cdot\,,\cdot)$ that maps pairs of vectors from V to real numbers that satisfies the following properties.

1. $(\mathbf{u},\mathbf{v}) = (\mathbf{v},\mathbf{u})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V.

2. $(\alpha\ \mathbf{u}+\beta\ \mathbf{v},\mathbf{w}) = \alpha(\mathbf{u},\mathbf{w}) + \beta(\mathbf{v},\mathbf{w})$ and $(\mathbf{w},\alpha\mathbf{u}+\beta\mathbf{v}) = \alpha(\mathbf{w},\mathbf{u}) + \beta(\mathbf{w},\mathbf{v})$ for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all real numbers α and β.

3. $(\mathbf{u},\mathbf{u}) \geq 0$ for all vectors $\mathbf{u} \in V$ and $(\mathbf{u},\mathbf{u}) = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

Definition Two vectors $\mathbf{u}$ and $\mathbf{v}$ in a vector space are said to be *orthogonal* with respect to an inner product if $(\mathbf{u},\mathbf{v}) = 0$.

Examples

The standard inner product on $\mathbb{R}^n$ is the vector dot product.

$$(\mathbf{u},\mathbf{v}) = \sum_{i=1}^{n} \mathbf{u}_i\, \mathbf{v}_i$$

The standard norm on $\mathbb{R}^n$ is

$$||\mathbf{u}|| = \sqrt{(\mathbf{u},\mathbf{u})}$$

The vector space $C[0,1]$ of continuous functions on the interval [0,1] has an inner product

$$(f,g) = \int_0^1 f(x)\ g(x)\, dx$$

This inner product is known as the L^2 inner product. Likewise, we can define an L^2 norm for this vector space by

$$||f\,|| = \sqrt{\int_0^1 (f(x))^2\, dx}$$

Other possible norms include the L^1 norm

$$||f|| = \int_0^1 |f(x)|\, dx$$

and the L^∞ norm

$$||\, f\,|| = max_{x \in [0,1]}\, |f(x)|$$

Orthogonal Bases

Definition A basis $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ for a vector space V is an orthonormal basis if $(\mathbf{v}_i\, ,\mathbf{v}_j) = 0$ for all $i \neq j$ and $(\mathbf{v}_i\, ,\mathbf{v}_i) = 1$ for all i.

Observation If a vector space has an orthonormal basis, computing coordinate representations with respect to that basis is very easy. Given an arbitrary vector $\mathbf{v}$ in V, we seek to compute a coordinate vector $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$ such that

$$c_1\, \mathbf{v}_1 + c_2\, \mathbf{v}_2 + \cdots + c_n\, \mathbf{v}_n = \mathbf{v}$$

If the basis is orthonormal, we can easily compute the coordinates c_i by taking the inner product with respect to v_i on both sides of the equation:

$$(c_1\, \mathbf{v}_1 + c_2\, \mathbf{v}_2 + \cdots + c_n\, \mathbf{v}_n\, ,\, v_i) = (\mathbf{v},\mathbf{v}_i)$$

$$c_1\, (\mathbf{v}_1,\mathbf{v}_i) + c_2\, (\mathbf{v}_2,\mathbf{v}_i) + \cdots + c_n\, (\mathbf{v}_n,\mathbf{v}_i) = (\mathbf{v},\mathbf{v}_i)$$

$$c_1\, 0 + c_2\, 0 + \cdots + c_i\, 1 + \cdots + c_n\, 0 = (\mathbf{v},\mathbf{v}_i)$$

$$c_i = (\mathbf{v},\mathbf{v}_i)$$

Constructing an Orthonormal Basis

The Gram-Schmidt algorithm is an algorithm that can convert a basis for a vector space into an alternative basis that is orthonormal. Here is an outline of that algorithm. Let $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ be a basis for a vector space V.

1. Convert the vector $\mathbf{v}_1$ into a normal vector by dividing it by its own norm.

$$\mathbf{u}_1 = \frac{1}{||\mathbf{v}_1||}\, \mathbf{v}_1$$

2. Construct

$$\mathbf{p}_2 = \mathbf{v}_2 - (\mathbf{u}_1\, ,\, \mathbf{v}_2)\, \mathbf{u}_1$$

The term $(\mathbf{u}_1 , \mathbf{v}_2)\, \mathbf{u}_1$ is the *projection* of $\mathbf{v}_2$ onto $\mathbf{u}_1$. By construction, $\mathbf{p}_2$ is orthogonal to $\mathbf{v}_1$ (why?).

3. We then form

$$\mathbf{u}_2 = \frac{1}{||\mathbf{p}_2||}\, \mathbf{p}_2$$

in order to make $\mathbf{u}_2$ be both normal and orthogonal to $\mathbf{u}_1$.

4. Next, compute

$$\mathbf{p}_3 = \mathbf{v}_3 - (\mathbf{u}_1 , \mathbf{v}_3)\, \mathbf{u}_1 - (\mathbf{u}_2 , \mathbf{v}_3)\, \mathbf{u}_2$$

and subsequently

$$\mathbf{u}_3 = \frac{1}{||\mathbf{p}_3||}\, \mathbf{p}_3$$

to produce a vector that is normal and perpendicular to both $\mathbf{u}_1$ and $\mathbf{u}_2$.

5. The process repeats until all of the original $\mathbf{v}_i$ vectors have been processed. The result is a set of $\mathbf{u}_i$ vectors which form an orthonormal basis for V.

The Projection Theorem

Here is a theorem from the text which also makes use of the concept of a projection.

Projection Theorem Let V be a vector space with an inner product. Let W be a finite dimensional subspace of V and let $\mathbf{v}$ by an arbitrary vector in V.

1. There is a unique $\mathbf{u}$ in W such that

$$||\mathbf{v} - \mathbf{u}|| = min_{\mathbf{w} \in W} ||\mathbf{v} - \mathbf{w}||$$

$\mathbf{u}$ is known as the *projection of* $\mathbf{v}$ *onto the subspace* W.

2. $(\mathbf{v}-\mathbf{u},\mathbf{z}) = 0$ for all $\mathbf{z} \in W$.

3. If $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ is a basis for W then

$$\mathbf{u} = \sum_{i=1}^{n} x_i\, \mathbf{w}_i$$

where

$$G\, \mathbf{x} = \mathbf{b}$$

$$G_{i,j} = (\mathbf{w}_i , \mathbf{w}_j)$$

$$\mathbf{b}_i = (\mathbf{w}_i , \mathbf{v})$$

The matrix G is known as the *Gram matrix* and the equations $G\, \mathbf{x} = \mathbf{b}$ are known as the *normal* *equations.*

4. If $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n\}$ is an orthogonal basis for W then

$$u = \sum_{i=1}^{n} \frac{(w_i\ ,\ v)}{(w_i\ ,\ w_i)}\ w_i$$

Observation A very important thing to note about the projection theorem is that the original vector space V does not have to be a finite dimensional space. The only requirement in the theorem is that W must be a finite dimensional subspace of V.

This opens an intriguing possibility. Suppose we have a linear operator f that maps V to V. If we want to make a finite representation for f we might do the following:

1. For a $\mathbf{v} \in V$ we compute the projection $\mathbf{u}$ of $\mathbf{v}$ onto W.

2. We compute $f(\mathbf{u})$ and hope that $f(\mathbf{u})$ stays in W. If it does not, we project $f(\mathbf{u})$ back onto the subspace W to make a vector $\mathbf{y}$.

3. What we have constructed is a r*estriction* of the operator f onto the subspace W. If f is still linear on W, we can construct a finite representation for the restricted operator and eventually represent that as a matrix A such that

$$A\ \mathbf{u} = \mathbf{y}$$