5374 Fall '22

Numerical Linear Algebra Lecture

Problem Set 2 Erratum:

In 4c), "$\|A\| \geq 1$ and "$\|A^{-1}\| \geq 1$" should be "$\|A\| \leq 1$ and "$\|A^{-1}\| \leq 1$"

More on inner products and orthogonality

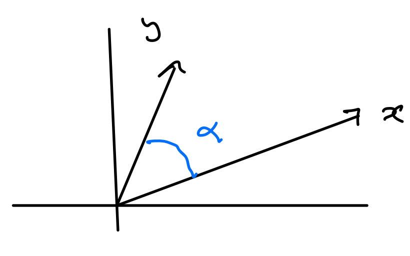
The Cauchy-Schwartz inequality says that

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|$$

where its understood that $\|\cdot\|$ is the norm induced by the inner product.

In the case of $\mathbb{R}^2$ with the usual inner product a basic fact about the inner product is the following:

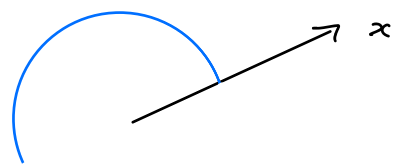
If $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$



$$x_1 y_1 + x_2 y_2 = \sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}\cos(\alpha)$$

In particular,

$$\alpha = \arccos\left(\frac{x_1 y_1 + x_2 \cdot y_2}{\sqrt{x_1^2+x_2^2}\sqrt{y_1^2+y_2^2}}\right) \in [0, \pi]$$



Now generalizing this fact from plane geometry, given any vector space V with an inner product we define the angle between two vectors $x, y \in V$ as

$$\alpha_{x,y} = \arccos\left(\frac{(x,y)}{\|x\| \cdot \|y\|}\right)$$

Thanks to the Cauchy-Schwartz inequality this $\alpha_{x,y}$ is well defined

With an inner product we can now talk not just about distances but also about angles. In particular x and y have a right angle between them when $(x,y)=0$, in this case we say x and y are orthogonal and denote it by $x \perp y$.

A family of vectors $q_1, \dots, q_k$ is called orthonormal if they are all of length 1 and are also pairwise orthogonal, i.e.

$$(q_i, q_j) = 0 \quad \text{if} \quad i \neq j$$
$$(q_i, q_i) = 1 \quad \text{for each } i$$

Pythagorean theorem

Suppose $\bar{x}_1, \dots, \bar{x}_k$ are pairwise orthogonal. Then

$$\| \bar{x}_1 + \dots + \bar{x}_k \|^2 = (\bar{x}_1 + \dots + \bar{x}_k, \bar{x}_1 + \dots + \bar{x}_k)$$

$$= (\bar{x}_1, \bar{x}_1 + \dots + \bar{x}_k) + (\bar{x}_2, \bar{x}_1 + \dots + \bar{x}_k) + \dots + (\bar{x}_n, \bar{x}_1 + \dots + \bar{x}_k)$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{k} (\bar{x}_i, \bar{x}_j)$$

Since $(\bar{x}_i, \bar{x}_j) = 0$ when $i \neq j$, the above sum simplifies to

$$\sum_{i=1}^{k} (\bar{x}_i, \bar{x}_i)$$

$$\| \bar{x}_1 + \cdots + \bar{x}_k \|^2 = \| \bar{x}_1 \|^2 + \cdots + \| \bar{x}_k \|^2$$

So this is the generalized Pythagorean theorem.

Similarly, for two vectors $\bar{x}_1, \bar{x}_2$ (not assumed to be orthogonal)

$$\begin{aligned} \| \bar{x}_1 + \bar{x}_2 \|_2^2 &= (\bar{x}_1 + \bar{x}_2, \bar{x}_1 + \bar{x}_2) \\ &= (\bar{x}_1, \bar{x}_1) + (\bar{x}_1, \bar{x}_2) + (\bar{x}_2, \bar{x}_1) + (\bar{x}_2, \bar{x}_2) \\ &= \| \bar{x}_1 \|^2 + \| \bar{x}_2 \|^2 + 2(\bar{x}_1, \bar{x}_2) \end{aligned}$$

(Law of cosines)

Proposition: If $\bar{x}_1, \ldots, \bar{x}_k$ are orthonormal, then they must be linearly independent.

For suppose $\exists$ numbers $c_1, \ldots, c_k$ s.t.

$$c_1 \bar{x}_1 + \cdots + c_k \bar{x}_k = 0$$

Then, $(c_1 \bar{x}_1 + \cdots + c_k \bar{x}_k, \bar{x}_i) = 0$

and the left hand side is equal to

$$c_i \cdot 1 = 0$$

This is true for every i, so $c_1 = \cdots = c_n = 0$
$\Rightarrow$ the vectors are linearly independent.

We will be especially concerned with bases that are also orthonormal.

If $\bar{q}_1, \ldots, \bar{q}_n$ is an orthonormal basis of V, and $\bar{x} \in V$. Then x can be expressed as

$$\bar{x} = (\bar{x}, \bar{q}_1)\bar{q}_1 + \cdots + (\bar{x}, \bar{q}_n)\bar{q}_n$$

The numbers $(\bar{x}, \bar{q}_i)$ $(i = 1, \ldots, n)$ are known as Fourier coefficients.

Example   The following formulas are well known:

Let $k_1, k_2 \in \mathbb{Z}$

$$\int_0^{2\pi} \cos(k_1 x)\cos(k_2 x)\,dx = 0 \qquad \text{if } k_1 \neq k_2$$

$$\int_0^{2\pi} \sin(k_1 x)\sin(k_2 x)\,dx = 0 \qquad \text{if } k_1 \neq k_2$$

$$\int_0^{2\pi} \sin(k_1 x)\cos(k_2 x)\,dx = 0$$

(Equivalently,

$$\int_0^{2\pi} e^{ik_1 x} \overline{e^{ik_2 x}} \, dx = \begin{cases} 0 & \text{if } k_1 \neq k_2 \\ 2\pi & \text{if } k_1 = k_2 \end{cases}$$

Since $e^{ik_1 x} = \cos(k_1 x) + i \sin(k_1 x)$ the formulas above all follow from this last one.)

Given a function $f : [0, 2\pi] \to \mathbb{R}$, then f can be expressed as

$$a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx)$$

where

$$a_k = \frac{\int_0^{2\pi} f(x) \cos(kx) \, dx}{\int_0^{2\pi} (\cos(kx))^2 \, dx}, \quad k = 0, 1, 2, \ldots$$

and so on for b_k ($k = 1, \ldots$)

Transpose and adjoint

Let V be a real vector space with an inner product.

Let $L : V \to V$ be a linear transformation.

There exists a unique linear transformation $L^t : V \to V$ such that

$$\forall\, x, y \in V \qquad (Lx, y) = (x, L^t y)$$

Let's see why this is so in $\mathbb{R}^n$ with its usual product:

$$\bar{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \bar{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

The inner product $(\bar{x}, \bar{y})$ is conveniently represented via matrix multiplication

$$\bar{y}^t \bar{x} = (y_1 \; \cdots \; y_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be linear, and let A be the matrix that represents L in the canonical basis $\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix}, \ldots \text{etc} \right)$, or what is the same, A is the unique matrix such that

$$L\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Then $\left(L(\bar{x}), \bar{y} \right) = \bar{y}^t (A \bar{x})$

But this is a matrix multiplication, and matrix multiplication is associative, so

$$\bar{y}^t (A\bar{x}) = (\bar{y}^t A)\bar{x}$$

On the other hand $(\bar{y}^t A)^t = A^t (\bar{y}^t)^t$
$= A^t \bar{y}$

This means $\bar{y}^t A = (A^t \bar{y})^t$, so

$$(L\bar{x}, \bar{y}) = \bar{y}^t (A\bar{x}) = (A^t y)^t \bar{x} = (\bar{x}, L^t(y))$$

where $L^t(y) := A^t y$.

This is the key property of matrix transpose

(For complex vector spaces, say $\mathbb{C}^n$, the relevant operation is the adjoint, if

$$A = (a_{ij})_{m \times n}$$

Then

$$A^* = (\bar{a}_{ji})_{n \times m}$$

Note $A^* = \overline{A^t}$)

This characterization of A^t is essential in proving facts about the transpose.

Definition: A matrix is called symmetric if

$$A = A^t$$

anti-symmetric if

$$A = -A^t$$

(also: Hermitian)

For complex matrices, A is called self-adjoint if

$$A = A^* .$$

Definition: A matrix is called positive semi-definite if $A = A^t$ and

$$(Ax, x) \geq 0 \quad \text{for all } x$$

Moreover, if $(Ax, x) > 0$ when $x \neq 0$, A is called positive definite.

Example $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ is positive definite

$B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is positive semi-definite, but not positive definite.

Example: Let A be any $m \times n$ matrix.

Let

$$M = A \cdot A^t$$

(this product makes sense for all values of $m \times n$)

Observe $\quad M^t = (A \cdot A^t)^t = (A^t)^t \cdot A^t = A \cdot A^t$

So $\quad M$ is symmetric.

Claim: $\quad M$ is always positive semidefinite.

Why?
$$
\begin{aligned}
&(Mx, x) \\
&= ((A \cdot A^t)x, x) \\
&= (A(A^t x), x) \\
&= (A^t x, A^t x) = \|A^t x\|^2 \geq 0
\end{aligned}
$$

Definition: A $n \times n$ matrix Q is called orthogonal if

$$Q^{-1} = Q^t.$$

Exercise: Show that the above holds if and only if the columns of Q form an orthonormal basis

Hint: Given A, B, both $n \times n$, the entries of AB represent inner products of rows and columns