Prove that under specific assumptions on P, $\sqrt{x^T P x}$ is a valid norm. Assume $x \in \mathbb{R}^n$ & $P \in \mathbb{R}^{n \times n}$

Proof: Suppose P is symmetric positive definite: i.e $P^T = P$ & $\forall x \neq 0 \; x^T P x > 0$

The condition $\forall\ \mathbf{x} \neq \mathbf{0}, \sum_{i=1} \sum_{j=1} a_{ij}\mathbf{x}_i\mathbf{x}_j > 0$ involves a quadratic expression. The expression is guaranteed to be greater than $0\ \forall\ \mathbf{x} \neq \mathbf{0}$ *iff* it can be expressed as $\sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{i-1} \beta_{ij} x_{ij} + x_{ii} \right)^2$, where $\lambda_i \geq 0$. This is possible *iff* A can be expressed as LDL^T, where, L is a lower traingular matrix with 1 in each diagonal entry and D is a diagonal matrix of all positive diagonal entries. Or equivalently, it should be possible to factorize A as RR^T, where $R = LD^{1/2}$ is a lower traingular matrix. Note that any symmetric matrix A can be expressed as LDL^T, where L is a lower traingular matrix with 1 in each diagonal entry and D is a diagonal matrix; positive definiteness has only an additional requirement that the diagonal entries of D be positive. This gives another equivalent condition for positive definiteness: *Matrix A is p.d. if and only if, A can be uniquely factored as $A = RR^T$, where R is a lower traingular matrix with positive diagonal entries.* This factorization of a p.d. matrix is reffered to as *Cholesky factorization.*

Source: pg 207 of http://www.cse.iitb.ac.in/~CS709/notes/LinearAlgebra.pdf

$$\Rightarrow x^T P x = x^T \underbrace{R R^T}_{\text{Assume } P = RR^T} x = \underbrace{(R^T x)^T}_{\text{Let } R^T x = y} (R^T x) = y^T y = \|y\|_2^2$$

$\therefore$ ① $x^T P x \geq 0$ since P is positive definite
$\&$ $x^T P x = 0$ iff $x = 0$ (By definition)

② $\|\alpha x\|_P = \sqrt{(\alpha x)^T P (\alpha x)} = \sqrt{\alpha^2 x^T P x}$
$= |\alpha| \, \|x\|_P$

③ $\|x+y\|_P^2 = (x+y)^T P (x+y) = (x+y)^T R R^T (x+y)$

$= x^T R \underbrace{R^T x}_{u} + y^T R \underbrace{R^T y}_{v} + x^T R R^T y$
$+ y^T R R^T x$

$= u^T u + v^T v + u^T v + v^T u$

$= \|u\|_2^2 + \|v\|_2^2 + 2u^T v$

$(\|x\|_P + \|y\|_P)^2 = \underbrace{\|x\|_P}^2 + \underbrace{\|y\|_P} + 2\|x\|_P \|y\|_P$
$= \|u\|_2^2 + \|v\|_2^2 + 2\sqrt{\|u\|_2^2 \|v\|_2^2}$

$= \|u\|_2^2 + \|v\|_2 + 2\|u\|_2 \|v\|_2$

Next we prove the Cauchy Schwarz inequality:
$2u^T v \leq \|u\|_2 \|v\|_2 \Rightarrow \|x+y\|_P \leq \|x\|_P + \|y\|_P$

In general (see http://en.wikipedia.org/wiki/Cauchy%E2%80%93Schwarz_inequality)

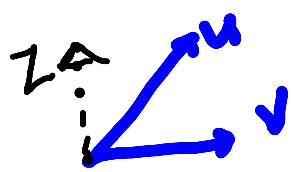
$$|\langle u,v\rangle| \leq \|u\|\,\|v\|$$ for any valid norm such as $\|\ \|_2$

Proof: If $v=0$, both sides are 0 & hence equality holds.

Assume $v \neq 0$ & let $z = u - \frac{\langle u,v\rangle}{\langle v,v\rangle} v$ {$z=0$ iff u & v are lin. dependent}

& $\therefore$ $\langle z,v\rangle = \left\langle u - \frac{\langle u,v\rangle}{\langle v,v\rangle} v, v\right\rangle = \langle u,v\rangle - \frac{\langle u,v\rangle}{\langle v,v\rangle}\langle v,v\rangle = 0$

(By linearity of the inner product in the first argument)



& $\therefore$ $\langle u,u\rangle = \|u\|^2 = \underbrace{\langle z,z\rangle + \left|\frac{\langle u,v\rangle}{\langle v,v\rangle}\right|^2 \langle v,v\rangle}_{} + \underbrace{\left|\frac{\langle u,v\rangle}{\langle v,v\rangle}\right| \langle z,v\rangle}_{=0 \text{ from above}}$

Substituting for $u = z + \frac{\langle u,v\rangle}{\langle v,v\rangle} v$

$$= \|z\|^2 + \left(\frac{\langle u,v\rangle^2}{\|v\|^2}\right) \geq \frac{\langle u,v\rangle^2}{\|v\|^2}$$ {equality iff $z=0$}

$\Rightarrow$ $\boxed{\|u\|\,\|v\| \geq |\langle u,v\rangle|}$ $\rightarrow$ Cauchy Shwarz ineq.

{equality iff u & v are linearly dependent}

[H/w: Prove that "inner product space" is a "normed" vector space]

Inner product space: It is a vector space over a field of scalars along with an inner product

Assume $\mathbb{R}$ or complex

① $\langle x, x\rangle = \overline{\langle x, x\rangle} \Rightarrow \langle x, x\rangle$ must be real

$\therefore$ We can define $\|x\| = \sqrt{\langle x, x\rangle}$

We need to prove that $\|x\|$ is a valid norm

ⓐ • By defn of inner product, since $\langle x, x\rangle \geq 0$ with equality iff $x = 0$, $\|x\| \geq 0$ iff $x = 0$

ⓑ • $\|tx\| = \sqrt{\langle tx, tx\rangle} = \sqrt{t\bar{t}\langle x, x\rangle}$

$= \sqrt{t\bar{t}}\,\|x\| = |t|\,\|x\|$ (For real & complex t, $|t| = \sqrt{t\bar{t}}$)

(c) $\|x+y\| = \sqrt{\langle x+y, x+y\rangle}$

$$= \sqrt{\langle x,x\rangle + \langle y,y\rangle + \langle x,y\rangle + \langle y,x\rangle}$$

$$\leq \sqrt{\langle x,x\rangle + \langle y,y\rangle + \sqrt{\langle x,x\rangle\langle y,y\rangle}\times 2}$$

By Cauchy Schwarz inequality

$$= \sqrt{(\|x\| + \|y\|)^2}$$

$$= \|x\| + \|y\|$$

– Hence proved that $\sqrt{\langle x,x\rangle}$ is a norm

$\Rightarrow$ Every inner product space is a normed space.

Converse does not hold: $\exists$ normed spaces that are not inner product spaces.

<u>Eg:</u> $\|x\|_p = \left[\sum_{i=1}^{\infty} |x_i|^p\right]^{1/p}$

Note a H/w problem for 7th August:

http://www.cse.iitb.ac.in/~cs709/calendar2013.html

- 31/07/2013. Show that the following are vector spaces (assuming scalars come from a set S), and then answer questions that follow for each of them: Set of all matrices on S, set of all polynomials on S, set of all sequences of elements of S. (HINT: You can refer to this book for answers to most questions in this homework.) How would you understand the concepts of independence, span, basis, dimension and null space (chapter 2 of this book), eigenvalues and eigenvectors (chapter 5), inner product and orthogonality (chapter 6)? EXTRA: Now how about set of all random variables and set of all functions. **Deadline:** August 7 2013.

Let us consider space of matrices:

$$\left\{ \begin{bmatrix} s_{11} & \dots & s_{1m} \\ & & \\ s_{n1} & \dots & s_{nm} \end{bmatrix} \middle| \; s_{11} \dots s_{nm} \in S \right\}$$ over scalars S

So far we considered $S = \mathbb{R}$

Obvious that this is a vector space (since multiplication etc are defined on S)

For simplicity, let $S = \mathbb{R}$ & let us consider a norms for matrices, induced by norms for vectors

Let $N(x)$ be a vector norm satisfying the vector norm axioms:

Then we will define a matrix norm

$$M_N(A) = \sup_{x \neq 0} \frac{N(Ax)}{N(x)}$$

Can you prove that this is indeed a valid norm?

$\sup f(s) = \hat{f}$

$s \in S$ if $\hat{f}$ is minimum upper bnd

as the matrix norm induced by vector norm $N(x)$

what, for example, will be

$M_N(I)$ → Ans: 1

irrespective of $N(x)$?

examples

$$\left(\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \right)$$

(a) If $N(x) = \|x\|_1 = \sum_{i=1}^{n} |x_i|$

$$\|A\|_1 = ?$$

Ans: $$\|Ax\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{m} a_{ij} x_j \right| \leq \sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}| |x_j|$$

abs value of sum $\leq$ sum of abs values

Changing order of summation:

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

→ Let $C = \max_j \sum_{i=1}^{n} |a_{ij}|$

Then $\|Ax\|_1 \le C\|x\|_1$

$$\Rightarrow \quad \|A\|_1 = \sup_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} \le C$$

But consider an $x = [0 \cdot 0 \cdot 1 \cdot \cdot 0\ 0]$

k^{th} position, where k is column index j for which $C = \sum_{i=1}^{n} |a_{ik}|$

Then $\|x\|_1 = 1$ & $\|Ax\|_1 = C$ (Show this)

$$\Rightarrow \quad \|A\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

i.e if $N(x) = \|x\|_1$ then $M_N(A) = \max_j \sum_i |a_{ij}|$

(b) Similarly,

if $N(x) = \|x\|_2 = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2}$

$$\|A\|_2 = [\text{dominant eigenvalue of } A^T A]^{1/2}$$

(c) If $N(x) = \|x\|_\infty = \max_i |x_i|$ $\left\{\lim_{p\to\infty} \left(\sum |x_i|^p\right)^{1/p}\right.$

$$\|A\|_\infty = \max_i \sum_{j=1}^{m} |a_{ij}|$$

Other matrix norms:

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$$

Frobenius norm

Q: What abt inner products:

Note: Not all normed spaces are inner prod spaces.

Eg: $\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$ → for $p=2$ $\langle x,y\rangle = \sum_i x_i y_i$

→ For $p=1$ or ∞, No corresp. inner products

Read more on

http://www.math.ucsd.edu/~njw/Teaching/Math271C/Lecture_04.pdf

Eg of Frobenius inner product:

$$\langle A,B\rangle = \sum_i \sum_j a_{ij} b_{ij}$$

weighted inner product

$$\langle A,B\rangle_w = \sum_i \sum_j a_{ij} b_{ij} w_{ij} \quad \text{for } w_{ij} > 0$$

Basis for vector space of matrices $(m \times n)$

$$\left\{ \underbrace{\underbrace{\begin{bmatrix} 1 & 0 & \cdot & 0 \\ \cdot & & & \\ \vdots & & & \\ 0 & \cdots & & 0 \end{bmatrix}}_{E_{11}} \underbrace{\begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & & \cdots \\ & & & \\ 0 & & \cdots & 0 \end{bmatrix}}_{E_{12}} \cdots}_{} \underbrace{\begin{bmatrix} 0 & \cdots & & 0 \\ \vdots & & & \\ 0 & \cdots & & 1 \end{bmatrix}}_{E_{mn}} \right\}$$

$m \times n$ linearly independent elements that span the space of all matrices of size $m \times n$

$$B = \sum_{i,j} a_{ij} E_{ij} \equiv \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{mn} \end{bmatrix}$$

This vector $\in \mathbb{R}^{mn}$ / S^{mn} is a canonical representation of B