Matrices and Matrix Norms.

We need to generalize the absolute value $|\ |$ to a measure of distance on higher-dimensional spaces and spaces of matrices.

So

Definition. Let $\mathbb{B}$ be a real or complex vector space ($\mathbb{R}^n$ or $\mathbb{C}^n$). It is <u>normed</u> if there's a function

$$\|\cdot\|: \mathbb{B} \to \mathbb{R}$$

so that

1) $\|\vec{x}\| \geq 0$ and $\|\vec{x}\| = 0 \iff \vec{x} = \vec{0}$
2) $\|\alpha\vec{x}\| = |\alpha|\ \|\vec{x}\|$ for any scalar α
3) $\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

The most common norms are

$$\|\vec{x}\|_p = \left(\sum_i |x_i|^p\right)^{1/p} \quad - \ell_p \text{ norm}$$

$$\|\vec{x}\|_\infty = \max |x_i| \quad - \ell_\infty \text{ norm}$$

Definition. A function $\langle \cdot , \cdot \rangle : \beta \to \mathbb{R}$ or $\langle \cdot , \cdot \rangle : \beta \to \mathbb{C}$ is an inner product if

1) $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$
2) $\langle \vec{x}, \vec{y}+\vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$
3) $\langle \alpha\vec{x}, \vec{y} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle$
4) $\langle \vec{x}, \vec{x} \rangle \geq 0$, $\langle \vec{x}, \vec{x} \rangle = 0 \iff \vec{x} = 0$

A standard inner product is

$$\langle \vec{x}, \vec{y} \rangle = \sum x_i \overline{y_i}$$

Lemma. For any inner product,

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}$$

Lemma. $\sqrt{\langle x,x\rangle}$ is a norm.

As we've seen before in our discussion of conjugate gradient, inner products are <u>quadratic forms</u>, which is to say

Lemma. Every inner product $\langle\, , \,\rangle$ can be written in terms of a symmetric (or Hermitian) positive definite matrix A as

$$\langle \vec{x}, \vec{y} \rangle = \vec{y}^* A \vec{x}.$$

Now all norms on $\mathbb{R}^n$ (or $\mathbb{C}^n$) are equivalent "up to constants" in the sense that $\exists$ constants $c_1, c_2 > 0$ so that

$$c_1 \|\vec{x}\|_a < \|\vec{x}\|_b < c_2 \|\vec{x}\|_a$$

for all $\vec{x}$.

Sometimes it's useful to know the constants:

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

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

$$\|\vec{x}\|_\infty \leq \|\vec{x}\|_1 \leq n\,\|\vec{x}\|_\infty$$

all of these are more-or-less obvious when you do them for homework

Definition. $\|\cdot\|$ is a <u>matrix norm</u> if it is a vector norm on $m \times n$ dimensional space:

$$\|A\| \geq 0 \;\vdots\; \|A\| = 0 \iff A = 0$$

$$\|\alpha A\| = |\alpha|\,\|A\|$$

$$\|A+B\| \leq \|A+B\|$$

$\max A_{ij} = \|\ \|_\infty$ is called the max norm, while $\| \; \|_2 = \left(\sum |a_{ij}|^2 \right)^{1/2}$ is called the <u>Frobenius</u> norm $\|A\|_F$.

We have a compatibility condition much like scalar multiplication "coming out" of a norm.

Definition. Norms on $m \times n$, $n \times p$ and $m \times p$ matrices are mutually consistent if

$$\|AB\| \leq \|A\| \, \|B\|$$

for any $m \times n$ matrix A and $n \times p$ matrix B.

Now the Frobenius norm thinks matrices are close if they have similar entries. Another strategy is to ~~g~~ regard matrices as close if they do similar things to vectors.

Definition. The <u>operator</u> <u>norm</u> on $m \times n$ matrices induced by norms on $\mathbb{R}^m$ and $\mathbb{R}^n$ is

$$\| A \| = \max_{\vec{x} \neq \vec{0}} \frac{\| A\vec{x} \|}{\| \vec{x} \|}$$

Lemma. An operator norm is a matrix norm.

We now ~~[illegible]~~ recall another definition:

Definition. Q is <u>orthogonal</u> if $Q^T = Q^{-1}$ and <u>unitary</u> if $Q^* = Q^{-1}$. (recall Q^* = conjugate transpose)

We now have a laundry list of useful norm properties. All of these are more or less easy to prove.

Lemma.

1) For any operator norm and corresponding vector norm (or the Frobenius norm and the 2-norm)

$$\|A\vec{x}\| \le \|A\| \, \|\vec{x}\|$$

2) For any operator norm or the Frobenius norm

$$\|AB\| \le \|A\| \, \|B\|$$

3) The max norm and Frobenius norm are not operator norms (for any norm on $\mathbb{R}^n$).

4) If Q, Z are orthogonal or unitary,

$$\|QAZ\| = \|Q\|$$

for the Frobenius norm and the operator norm induced by the 2-norm.

We now consider the operator norms induced by our favorite norms on $\mathbb{R}^n$: the ∞-norm, the 2-norm, and the 1-norm.

Lemma.

$$\|A\|_\infty = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|_\infty}{\|\vec{x}\|_\infty} = \max_i \sum_j |a_{ij}|$$

$$= \text{maximum absolute } \underline{\text{row}} \text{ sum}$$

$$\|A\|_1 = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|_1}{\|\vec{x}\|_1} = \max_j \sum_i |a_{ij}|$$

$$= \text{maximum absolute } \underline{\text{column}} \text{ sum}$$

$$= \|A^T\|_\infty$$

$$\|A\|_2 = \max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|_2}{\|\cancel{A}\vec{x}\|_2} = \sqrt{\lambda_{max}(A^*A)}$$

where λ_{max} is the largest eigenvalue of the square matrix A^*A.

$$= \|A^T\|_2.$$

If A is a square matrix and $A^*A = AA^*$, then

$$\|A\|_2 = \lambda_{max}(A).$$

As twith the vector norms, these operator norms are related by constants, so they aren't <u>that</u> different. If we plug in the constants from before...

Lemma. If A is nxn, and $\|A\|_F$ is the Frobenius norm:

$$n^{-1/2}\|A\|_2 \leq \|A\|_1 \leq n^{1/2}\|A\|_2$$

$$n^{-1/2}\|A\|_2 \leq \|A\|_\infty \leq n^{1/2}\|A\|_2$$

$$n^{-1}\|A\|_\infty \leq \|A\|_1 \leq n\|A\|_\infty$$

$$\|A\|_1 \leq \|A\|_F \leq n^{1/2}\|A\|_2$$