Math128B
January 28, 2005
Jonathan Dorfman
Background Material on Norms

1. $\|x\|_p$ denotes the p-norm of a vector in $\mathbb{R}^n$. If no p is indicated, then $p = 2$ is assumed.

2. We define the ‘operator p-norm’ of A as

$$\|A\|_p \overset{\text{def}}{=} \sup_{x \neq 0} \frac{\|A\|_p}{\|x\|_p} = \sup_{\|x\|_p = 1} \|Ax\|_p$$

3. (Medium)

$$< y,\ x > \overset{\text{def}}{=} y'x = \|x\|\ \|y\|\ \cos\angle(x, y)$$

An easy consequence of this is the ”Cauchy-Schwarz-Bunyakovsky Inequality” (with equality iff x and y parallel):

$$|< y,\ x >| \leq \|x\|\ \|y\|$$

4. (Medium) If $l : \mathbb{R}^n \to \mathbb{R}^1$ is linear then there is a unique $y \in \mathbb{R}^n$ such that $l(x) =< y,\ x > \forall x \in \mathbb{R}^n$. Since the $y \in \mathbb{R}^n$ depends on l, we write $y = y_l$. One should think of y_l as a row vector which operates on $x \in \mathbb{R}^n$ by dot-ing x with y_l:

$$l(x) = [\dots y_l \dots] \begin{bmatrix} \vdots \\ x \\ \vdots \end{bmatrix} \qquad \text{(clearly linear in } x\text{)}$$

5. (Easy) By (1), it makes sense to ask about the ‘operator norm’ of the linear operator $l : \mathbb{R}^n \to \mathbb{R}^1$. By (3) this turns out to be equal to $\|y_l\|$ where y_l is determined by (4).

6. (Hard) The operator norm defined in (2) can also be given by the following when $p = 2$:

$$\|A\| = \sup_{\|x\| = \|y\| = 1} < Ax, y >$$

7. (Easy) Using (5) and (6) we can prove $\|A^*\| = \|A\|$

8. (Easy) $\|AB\| \leq \|A\|\ \|B\|$. For general A and B equality is *not* achieved, (e.g. $A = B = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$)

9. (Easy) Using (7) and (8) we can prove $\|A^*A\| \leq \|A\|^2$

10. (Hard) Using (3) and (6) we can prove $\|A^*A\| \geq \|A\|^2$.

11. (Easy) (9) and (10) imply $\|A^*A\| = \|A\|^2$

12. (Easy) Using 11 we can prove if A is Hermitian, then $\|A^2\| = \|A\|^2$ (note that $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ example in (8) is not Hermitian).

13. For any $A : \mathbb{R}^n \to \mathbb{R}^n$ we define the spectrum of A:

$$\sigma(A) \overset{\text{def}}{=} \left\{ \lambda \in \mathbb{C}^n \middle| \quad \lambda \quad \text{is an eigenvalue of } A\text{, i.e. is a root of } A's \text{ characteristic polynomial} \right\}$$

14. For any $A : \mathbb{R}^n \to \mathbb{R}^n$ we define the spectral radius of A:

$$\rho(A) \overset{\text{def}}{=} \text{the maximum (complex) norm of the eignevalues of } A = \text{"size" of } \sigma(A)$$

15. (Easy) $\sigma(A^*) = \overline{\sigma(A)}$ and $\rho(A^*) = \rho(A)$ and $\rho(A^n) = \rho(A)^n$

16. (Easy) $\rho(A) \leq \|A\|$ ($\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ is example of strict inequality)

17. (Homework) If A is Hermitian, then $\rho(A) = \|A\|$

18. (Easy) (17) implies $\rho(A^*A) = \|A^*A\|$

19. (Easy) (18) and (11) imply $\rho(A^*A) = \|A\|^2$, i.e. $\|A\| = \sqrt{\rho(A^*A)}$ = formula for $\|A\|_2$

Some more definitions:

DEFINING PROPERTY	NAME	SPECTRUM	INTERPRETATION
$A^tA = AA^t = I$ real	orthogonal	$\sigma(A) \subset \left\{ z \in \mathbb{C} \middle\| \|z\| = 1 \right\}$	preserves Euclidean length
$A^*A = AA^* = I$	unitary	$\sigma(A) \subset \left\{ z \in \mathbb{C} \middle\| \|z\| = 1 \right\}$	complex version
$A^t = A$ real	symmetric	$\sigma(A) \subset \mathbb{R}$	stretch in perpendicular directions
$A^* = A$	Hermitian	$\sigma(A) \subset \mathbb{R}$	complex version
$A = B^*B$	positive	$\sigma(A) \subset \mathbb{R}^{\geq 0}$	symmetric, positive stretches
$A^{-1} = A$ $(A^2 = I)$	involution	$\sigma(A) \subset \left\{ -1, 1 \right\}$	oblique reflection
$A^{-1} = A = A^*$	reflection	$\sigma(A) \subset \left\{ -1, 1 \right\}$	perpendicular reflection
$A^2 = A$	idempotent	$\sigma(A) \subset \left\{ 0, 1 \right\}$	oblique projection
$A^2 = A = A^*$	projection	$\sigma(A) \subset \left\{ 0, 1 \right\}$	orthogonal projection
$A^*A = AA^*$	normal	$\sigma(A)$	iff diagonalizable (via unitary change of basis)

Additional facts:

1. Diagonalizable matrices share eigenspaces iff they commute iff simultaneously diagonalizable
2. Every real matrix is sum of symmetric and skew-symmetric ($A^t = -A$)
3. Every complex matrix is sum of Hermitian and skew-Hermitian ($A^* = -A$)
4. A matrix is (unitarily) diagonalizable iff its Hermitian and skew-Hermitian parts commute