The condition numbers

Measure how output changes when make a small change in the input.

View problem as a function $f(X)$

let δX be small change, and

$$\delta f = f(X+\delta X) - f(X)$$

Absolute condition number at x is

$$\hat{K} = \lim_{\varepsilon \to 0} \sup_{\|\delta x\| \le \varepsilon} \frac{\|\delta f\|}{\|\delta x\|}$$

Restrict changes to be small.

View δX as measurement error - error in input (but also comes from floating point)

So, we want $f(X)$, but get $f(X+\delta X)$

From this perspective $\frac{\|\partial f\|}{\|\partial x\|} = \frac{\text{forward error}}{\text{backward error}}$

If condition number is large at X, then we better be sure we know X to high accuracy.

Will call such X "ill-conditioned"

Note: restricting to small δX is reasonable, but feels arbitrary for some problems.

Ex. again consider invertible A,

and $f_A(b) = A^{-1}b$

So, $f_A(b+\delta b) = A^{-1}(b+\delta b)$ and

$$\delta f_A = f_A(b+\delta b) - f(b) = A^{-1}(\delta b)$$

$$\Rightarrow \frac{\|\delta f_A\|_2}{\|\delta b\|_2} = \frac{\|A^{-1}(\delta b)\|_2}{\|\delta b\|_2} \leq \|A^{-1}\|_2$$

Often consider the relative condition #:

$$K = \frac{\|f(\delta x)\| / \|f(x)\|}{\|\delta x\| / \|x\|}$$

For f_A, it equals $\hat{K} \cdot \frac{\|b\|}{\|A^{-1}b\|} = \hat{K} \frac{\|Ay\|}{\|y\|}$

where $y = A^{-1}b$.

So, $K \leq \|A^{-1}\| \cdot \|A\|$

Can find examples on which this is tight:

if δb satisfies $\|A^{-1}(\delta b)\| = \|A^{-1}\| \cdot \|\delta b\|$ and

$y = A^{-1}b$ satisfies $\|Ay\| = \|A\| \cdot \|y\|$.

Now, lets examine perturbations to A.

For simplicity we will fix b (for now).

Add δA to A, and look at change in solution, δy.

$$(A + \delta A)(y + \delta y) = b$$

Gives $Ay + A\delta y + \delta A y = b$, as $\delta A \delta y \to 0$ for small δA, δy

$$\Rightarrow A(\delta y) = -(\delta A) y \qquad (Ay = b)$$

$$\delta y = -A^{-1}(\delta A) y$$

$$\|\delta y\| \le \|\delta A\| \cdot \|A^{-1}\| \cdot \|y\|$$

$$\hat{K} = \frac{\|\delta y\|}{\|\delta A\|} \le \|A^{-1}\| \cdot \|y\|$$

$$K = \frac{\|\delta y\| / \|y\|}{\|\delta A\| / \|A\|} \le \|A\| \cdot \|A^{-1}\|$$

and this can be tight.

See Solomon or Demmel for an analysis in A and b together.

We define $K(A) = \|A\| \cdot \|A^{-1}\|$

Alternate approach:

For $f: \mathbb{R}^n \to \mathbb{R}^m$, $f(X+\delta X) \approx f(x) + \nabla f(x)^T \delta X$

But, what is ∇f for $f \to \mathbb{R}^m$?

It is the Jacobian: the matrix obtained by concatenating the gradients of each output of f

$$J = \left(\nabla f(x)(1), \ldots, \nabla f(x)(m) \right)$$

$$\hat{K} = \lim_{\varepsilon \to 0} \sup_{\|\delta X\|_2 \le \varepsilon} \frac{\|f(X+\delta X) - f(X)\|}{\|\delta X\|}$$

$$= \lim_{\varepsilon \to 0} \sup_{\|\delta X\|_2 \le \varepsilon} \frac{\|J \cdot \delta X\|_2}{\|\delta X\|_2} = \|J\|_2$$

Norms and the Singular Value Decomposition.

<u>Theorem</u> (proof in a later lecture)

Every square matrix A can be written $A = USV^T$,

where U and V are <u>orthogonal</u> and

S is nonnegative and diagonal.

The diagonal entries of $S(i,i) = \sigma_i$ are the <u>singular values</u>

Usually order so $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

Rectangular is OK, too. Just add zero rows or columns to make it square.

<u>Recall</u> U is orthogonal iff $U^T U = I$.

Called <u>orthogonal</u> because implies for

$i \neq j$, $0 = I(i,j) = (U^T U)(i,j) = U(:,i)^T U(:,j)$

So, columns of U are orthogonal.

And, $\|U(:,i)\|_2 = 1$ for all i.

Holds iff $UU^T = I$, because left-inverse = right-inverse for square matrices.

Key property of orthogonal matrices:

$\forall x \quad \|Ux\|_2 = \|x\|_2$

because $\|Ux\|_2^2 = x^T U^T U x = x^T x = \|x\|_2^2$

Formulas from the SVD:

$$A = USV^T \Rightarrow U^TAV = U^TUSV^TV = S$$

$$A^T = VS^TU^T = VSU^T$$

if $\sigma_n > 0$, $A^{-1} = (USV^T)^{-1} = V^{-T}S^{-1}U^{-1}$
$= V S^{-1} U^T$

and S^{-1} is diagonal with entries $1/\sigma_i$

let $u_1, \ldots, u_n$ be columns of U
and $v_1 \ldots v_n$ be columns of V.

$A = \sum_i \sigma_i u_i v_i^T$, a sum of rank-1 matrices

As $v_i^T v_j = 0$ for $i \neq j$, $Av_j = \sum_i \sigma_i u_i v_i^T v_j = \sigma_j u_j$

<u>Lem</u> $\|A\|_2 = \sigma_1$

<u>proof</u> $\|A\|_2 \geq \sigma_1$ because $\|Av_1\|_2 = \sigma_1 \|u_1\|_2 = \sigma_1$
(and $\|v_1\|_2 = 1 = \|u_1\|_2$)

For every x, $\|Sx\|_2 \leq \sigma_1 \cdot \|x\|_2$,
because $\|Sx\|_2^2 = \sum_i \sigma_i^2 x(i)^2 \leq \sum_i \sigma_1^2 x(i)^2 \leq \sigma_1^2 \|x\|^2$

So $\|Ax\|_2 = \|USV^Tx\|_2 = \|SV^Tx\|_2 \leq \sigma_1 \|V^Tx\|_2 = \sigma_1 \|x\|_2$

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

Cor. $\|A^{-1}\|_2 = 1/\sigma_n$. So $\kappa(A) = \dfrac{\sigma_1}{\sigma_n}$

Similarly, $\displaystyle\min_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sigma_n = \frac{1}{\|A^{-1}\|}$

Let P be the set of singular matrices.

Thm $\displaystyle\min_{A_0 \in P} \|A - A_0\|_2 = \sigma_n(A)$

proof 1. If $\|\Delta\|_2 < \sigma_n(A)$, then $A - \Delta \notin P$

because for all $\|x\|=1$, $\|Ax - \Delta x\|_2 \geq \|Ax\|_2 - \|\Delta x\|_2$

$\geq \sigma_n(A) - \|\Delta\|_2 > 0$

So is no x s.t. $(A-\Delta)x = 0$.

2. Let $\Delta = \sigma_n u_n v_n^T$

Then $A - \Delta = \displaystyle\sum_{i=1}^{n-1} \sigma_i u_i v_i^T$ has rank $\leq n-1$.

In particular, $(A-\Delta)v_n = 0$

And, $\|\Delta\|_2 = \sigma_n$.

The normalized distance, $\dfrac{dist(A, P)}{\|A\|_2} = \dfrac{\sigma_n(A)}{\sigma_1(A)} = \dfrac{1}{\kappa(A)}$

Also holds if measure distance in $\|A\|_F$

$$\left(\text{recall } \|A\|_F = \left(\sum A(i,j)^2\right)^{1/2}\right)$$

Lem $\|A\|_F = \left(\sum_i \sigma_i^2\right)^{1/2}$

proof First observe $\|A\|_F = \|U^T A\|_F$,

because $\|A\|_F^2 = \sum_i \|A(:,i)\|_2^2 = \sum_i \|U^T A(:,i)\|_2^2$

$$= \|U^T A\|_F^2$$

Similarly, $\|A\|_F = \|U^T A V\|_F = \|S\|_F = \left(\sum_i \sigma_i^2\right)^{1/2}$

Now, $\min_{A_0 \in P} \|A - A_0\|_F = \sigma_n$,

because $\|\Delta\|_2 \leq \|\Delta\|_F$, so part 1 holds.

And for rank-1 matrices like $\sigma_n u_n v_n^T$,

$$\|\sigma_n u_n v_n^T\|_2 = \|\sigma_n u_n v_n^T\|_F = \sigma_n$$