3/7/2019 Norms continued : Matrix Norms

Matrix Norms:

Definition 2.10 Given any norm $\|\cdot\|$ on $\mathbb{R}^n$, the subordinate matrix norm on the space $\mathbb{R}^{n\times n}$ is defined as:

$$\|\underline{\underline{A}}\| = \max_{\substack{\underline{v}\in\mathbb{R}^n \\ \underline{v}\neq\underline{0}}} \frac{\|\underline{\underline{A}}\,\underline{v}\|}{\|\underline{v}\|}.$$

Example: let $\underline{\underline{I}}\in\mathbb{R}^{n\times n}$ be the $n\times n$ identity matrix, and let $\|\cdot\|$ be any norm on $\mathbb{R}^n$. Then

$$\|\underline{\underline{I}}\| = \max_{\substack{\underline{v}\in\mathbb{R}^n \\ \underline{v}\neq\underline{0}}} \frac{\|\underline{\underline{I}}\,\underline{v}\|}{\|\underline{v}\|} = \frac{\|\underline{v}\|}{\|\underline{v}\|} = 1.$$

Rmk: Sometimes the subordinate matrix norm is called a matrix norm induced by the vector norm.

Some facts about actually computing matrix norms which are subordinate to the 1, 2, and ∞ vector norms:

Thm 2.7: $\|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}|$.

Thm 2.8: $\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|$.

Thm 2.9: $\|A\|_2 = \max_{i=1,\dots,n} \lambda_i^{1/2}$,

where λ_i are the eigenvalues of A^TA.

Pf of Thm 2.9

Let λ_i be the eigenvalues of A^TA, and $\underline{w}_i$ be the eigenvectors. Assume the eigenvectors $\underline{w}_i$ are orthonormal.

Note that the eigenvalues of A^TA are real since A^TA is symmetric, and nonnegative since

$$A^TA\,\underline{w}_i = \lambda_i \underline{w}_i$$

$$\Leftrightarrow \quad \underline{w}_i^T A^TA\,\underline{w}_i = \lambda_i \underline{w}_i^T \underline{w}_i$$

$$\Leftrightarrow \quad \lambda_i = \frac{\underline{w}_i^T A^TA\,\underline{w}_i}{\underline{w}_i^T \underline{w}_i} = \frac{\|A\underline{w}_i\|_2}{\|\underline{w}_i\|_2} \geq 0.$$

Take an arbitrary vector $\underline{u} \in \mathbb{R}^n$, $\underline{u} \neq \underline{0}$.

Express: $\underline{u} = c_1 \underline{w}_1 + \ldots + c_n \underline{w}_n$.

Notice that

$$\|\underline{u}\|_2^2 = \underline{u}^T \underline{u} = \sum_{i=1}^{n} c_i^2. \quad (\text{Why?})$$

Also, $A^T A \underline{u} = A^T A (c_1 \underline{w}_1 + \ldots + c_n \underline{w}_n)$

$$= \lambda_1 c_1 \underline{w}_1 + \ldots + \lambda_n c_n \underline{w}_n.$$

Without loss of generality, assume that

$$0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n.$$

Now, we have:

$$\|A\underline{u}\|_2^2 = \underline{u}^T A^T A \underline{u} = \underline{u}^T (\lambda_1 c_1 \underline{w}_1 + \ldots + \lambda_n c_n \underline{w}_n)$$

$$\left(\text{By } \underline{w}_i^T \underline{w}_j = 0 \text{ if } i \neq j\right) \quad = \lambda_1 c_1^2 + \ldots + \lambda_n c_n^2$$

$$\le (c_1^2 + \ldots + c_n^2) \lambda_n$$

$$= \lambda_n \|\underline{u}\|_2^2.$$

So, we have shown that

$$\frac{\|A\underline{u}\|_2^2}{\|\underline{u}\|_2^2} \le \lambda_n.$$

To see that equality is obtained, take $\underline{u} = \underline{w}_n$.

Another Fact:

Thm: 2.10 Given that $\|\cdot\|$ is a subordinate matrix norm,

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

Pf: By definition of subordinate matrix norm:

$$\|AB\| = \max_{\substack{\underline{v} \in \mathbb{R}^n \\ \underline{v} \ne \underline{0}}} \frac{\|AB\underline{v}\|}{\|\underline{v}\|}$$

Note that

$$\|AB\underline{v}\| = \frac{\|AB\underline{v}\|}{\|B\underline{v}\|} \|B\underline{v}\|$$

$$\le \max_{\substack{\underline{w} \in \mathbb{R}^n \\ \underline{w} \ne \underline{0}}} \frac{\|A\underline{w}\|}{\|\underline{w}\|} \, \|B\underline{v}\|$$

$$= \|A\| \; \|B\underline{v}\|,$$

i.e. we have $\|AB\underline{v}\| \le \|A\| \; \|B\underline{v}\|$.

Multiply by $\dfrac{1}{\|\underline{v}\|}$ for $\underline{v} \neq \underline{0}$ to get

$$\frac{\|AB\underline{v}\|}{\|\underline{v}\|} \le \|A\| \; \frac{\|B\underline{v}\|}{\|\underline{v}\|}.$$

Take max on both sides over $\underline{v} \in \mathbb{R}^n$, $\underline{v} \neq \underline{0}$, to get

$$\|AB\| \le \|A\| \; \|B\|.$$

Definition: $f: \mathcal{D} \subset V \to W$.

where $\|\cdot\|_V$ is a norm on V and $\|\cdot\|_W$ is a norm on W.

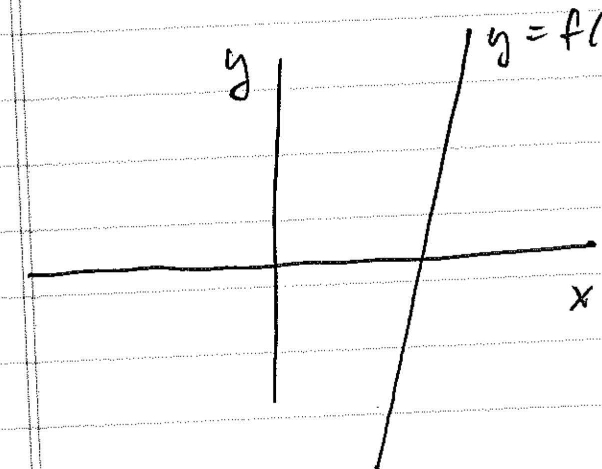
Absolute condition number of f

is defined to be:

$$\mathrm{Cond}(f) = \sup_{\substack{\underline{x}, \underline{y} \in \mathcal{D} \\ \underline{x} \neq \underline{y}}} \frac{\| f(\underline{y}) - f(\underline{x}) \|_W}{\| \underline{y} - \underline{x} \|_V}.$$

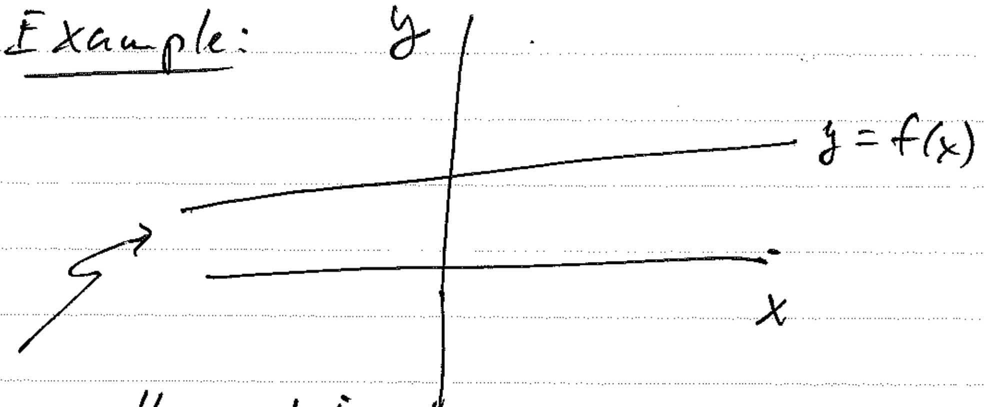
If $\mathrm{Cond}(f)$ is really really big, f is said to be *ill-conditioned*.

Example:



poorly, or ill-conditioned. Notice that small changes in x result in very large changes in $f(x)$.

Example:



This is well-conditioned, since small changes in x result in small changes in $f(x)$.

Definition: Absolute local condition number.

$$\mathrm{Cond}_x(f) = \sup_{\substack{\underline{\delta x} \in V \sim \{0\} \\ \underline{x} + \underline{\delta x} \in D}} \frac{\| f(\underline{x} + \underline{\delta x}) - f(\underline{x}) \|_W}{\| \underline{\delta x} \|_V}.$$

Definition: Relative local condition number.

$$\mathrm{cond}_x(f) = \sup_{\substack{\underline{\delta x} \in V \sim \{0\} \\ \underline{x} + \underline{\delta x} \in D}} \frac{\| f(\underline{x} + \underline{\delta x}) - f(\underline{x}) \|_W / \| f(\underline{x}) \|_W}{\| \underline{\delta x} \|_V / \| \underline{x} \|_V}.$$

Let's apply this idea of condition number to solving a linear system:

$$\underline{\underline{A}}\,\underline{x} = \underline{b}.$$

let $\underline{b} \neq \underline{0}$. You can think about $\underline{\underline{A}}^{-1}$ as a map from $\underline{b}$ to $\underline{x} = \underline{\underline{A}}^{-1}\underline{b}$, i.e. for a given right hand side $\underline{b}$, we "compute" $\underline{x} = \underline{\underline{A}}^{-1}\underline{b}$.

$$A^{-1}\cdot : \underline{b} \in \mathbb{R}^n \longmapsto \underline{\underline{A}}^{-1}\underline{b} \in \mathbb{R}^n.$$

let's compute the relative local condition number of $\underline{\underline{A}}^{-1}$ at $\underline{b}$:

$$\mathrm{cond}_{\underline{b}}(\underline{\underline{A}}^{-1}\cdot) = \sup_{\substack{\underline{\delta b} \in \mathbb{R}^n \\ \underline{\delta b} \neq \underline{0}}} \frac{\|A^{-1}(\underline{b}+\underline{\delta b}) - A^{-1}\underline{b}\| / \|A^{-1}\underline{b}\|}{\|\underline{\delta b}\| / \|\underline{b}\|}$$

$$= \|A^{-1}\| \frac{\|\underline{b}\|}{\|A^{-1}\underline{b}\|}.$$

We can write:

$$\|\underline{b}\| = \|A(A^{-1}\underline{b})\| \le \|A\| \, \|A^{-1}\underline{b}\|.$$

Plugging in this inequality, we have the statement:

$$\operatorname{cond}_{\underline{b}}(A^{-1}\cdot) \le \|A^{-1}\| \, \|A\|.$$

Also, for the mapping

$$A\cdot : \underline{x} \longmapsto A\underline{x},$$

you can show:

$$\operatorname{cond}_{\underline{x}}(A\cdot) \le \|A\| \, \|A^{-1}\|.$$

<u>Definition</u>: The condition number of a nonsingular matrix A is

$$\kappa(A) = \|A\| \, \|A^{-1}\|.$$