Honors Numerical Analysis

$\boxed{\text{Lecture 6}}$

Conditioning for matrix equations

Consider solving $A\vec{x} = \vec{b}$. The input is $\vec{b}$, and the output is $\vec{x} = A^{-1}\vec{b}$.

Let $\|\vec{b} - \vec{b}'\|$ be small, and $\vec{x} = A^{-1}\vec{b}$, $\vec{x}' = A^{-1}\vec{b}'$,

then
$$\|\vec{x} - \vec{x}'\| = \|A^{-1}\vec{b} - A^{-1}\vec{b}'\|$$

$$(*) \qquad \leq \underbrace{\|A^{-1}\|}_{\text{Absolute condition number of } A} \cdot \|\vec{b} - \vec{b}'\| \qquad \leftarrow \text{induced matrix norm}$$

But recall: The absolute condition number tells us nothing about the number of correct digits in the answer, for this we need the relative cond. number:

The previous expression $(*)$ implies that

$$\frac{\|\vec{x} - \vec{x}'\|}{\|\vec{x}\|} \leq \|A^{-1}\| \frac{\|\vec{b} - \vec{b}'\|}{\|\vec{x}\|}$$

$$= \|A^{-1}\| \frac{\|\vec{b} - \vec{b}'\|}{\|\vec{b}\|} \frac{\|\vec{b}\|}{\|\vec{x}\|} \qquad \frac{\|A\vec{x}\|}{\|\vec{x}\|} \leq \|A\|$$

$$\leq \underbrace{\|A\| \, \|A^{-1}\|}_{\text{Relative condition number of } A} \frac{\|\vec{b} - \vec{b}'\|}{\|\vec{b}\|}$$

(max × min stretching factors)

Now that we've discussed vector/matrix norms,

back to multivariate Newton... solve

$$\begin{aligned} f_1(x_1,\dots,x_n) &= 0 \\ &\vdots \\ f_n(x_1,\dots,x_n) &= 0 \end{aligned}$$

Multivariable Taylor Series

Let f be a scalar-valued function of n variables:

$$f = f(x_1, x_2, \dots, x_n)$$
$$= f(\vec{x})$$

Recall the Taylor series expansion for a function f of one variable:

$$f(x) = f(y) + f'(y)(x-y) + \frac{f''(y)}{2!}(x-y)^2 + \dots$$

$$= \sum_{j=1}^{\infty} \frac{f^{(j)}(y)}{j!}(x-y)^j$$

assuming f is C^∞ in a neighborhood of y.

Radius of convergence?

If $f = f(x_1,\dots,x_n)$, then we must incorporate partial derivatives into the expansion:

$$f(x_1,\dots,x_n) = \overbrace{f(\vec{y}) + \frac{\partial f}{\partial x_1}(\vec{y})(x_1-y_1) + \frac{\partial f}{\partial x_2}(\vec{y})(x_2-y_2) + \dots}^{\text{Linear approximation}}$$

$$\underbrace{+ \frac{1}{2!}\frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{y})(x_1-y_1)(x_2-y_2) + \frac{1}{2!}\frac{\partial^2 f}{\partial x_1^2}(\vec{y})(x_1-y_1)^2 + \dots}_{\text{Quadratic terms.}}$$

The general formula is slightly more complicated, but can be written out as:

$$f(\vec{x}) = f(\vec{y}) + \sum_{l=1}^{n} \frac{\partial f}{\partial x_l}(\vec{y})\,(x_l - y_l) + \frac{1}{2!}\sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_l}(\vec{y})(x_k - y_k)(x_l - y_l)$$

$$+ \underbrace{\frac{1}{3!}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n} \frac{\partial^3 f}{\partial x_k \partial x_l \partial x_m}(\vec{y})(x_k - y_k)(x_l - y_l)(x_m - y_m)}_{3^{rd}\text{-order terms}} + \dots$$

Multivariable vector-valued Taylor series

Now, consider a vector-field $\vec{f} = \vec{f}(x_1, .., x_n) = \vec{f}(\vec{x}) \in \mathbb{R}^n$.

$$f(\vec{x}) = \begin{pmatrix} f_1(x_1, .., x_n) \\ f_2(x_1, .., x_n) \\ \vdots \\ f_n(x_1, .., x_n) \end{pmatrix}$$

We can expand each component f_i in a multivariable Taylor series:

$$f_i(\vec{x}) = f_i(\vec{y}) + \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k}(\vec{y})\,(x_k - y_k) + \dots$$

If we stack these expansions into the vector $\vec{f}$, then we can collect terms and write this in the form:

$$\vec{f}(\vec{x}) = \vec{f}(\vec{y}) + \underbrace{J(\vec{y})}_{\substack{\text{matrix} \\ \text{Jacobian}}}(\vec{x}-\vec{y}) + \frac{(\vec{x}-\vec{y})^T \, \overbrace{Q(\vec{y})}}{2!}(\vec{x}-\vec{y}) + \ldots$$

$$\underbrace{Q(\vec{y})}_{\text{TENSOR}}$$

$$J(\vec{y}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \ddots & & \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_n} & & & \frac{\partial f_n}{\partial x_n} \end{pmatrix}(\vec{y})$$

$Q(\vec{y})$ is defined so that the i^{th}-entry of

$\frac{(\vec{x}-\vec{y})^T Q(\vec{y})(\vec{x}-\vec{y})}{2!}$ is given by $\underbrace{\frac{1}{2!}\sum_{k,l} \frac{\partial^2 f_i}{\partial x_k \partial x_l}(x_k - y_k)(x_l - y_l)}$

$$= \frac{1}{2}(\vec{x}-\vec{y})^T \underbrace{H_i(\vec{y})}(\vec{x}-\vec{y})$$

this is an $n \times n$ matrix called the Hessian of f_i.

- Higher order terms in this expression become much more complicated, and require "multi-index" notation.

Multivariate Newton's Method

Recall: A system of nonlinear equations is of the form:

$$\left.\begin{array}{c} f_1(x_1,\ldots,x_n)=0 \\ \vdots \\ f_n(x_1,\ldots,x_n)=0 \end{array}\right\} \quad \vec{f}(\vec{x})=\vec{0}.$$

In general, determining whether this has one, none, or many solutions is exceedingly difficult.

If we approximate $\vec{f}$ at $\vec{x}_0$ near the root $\vec{\xi}$ (assuming that such a root exists) then we have that:

$$\vec{f}(\vec{x}) \approx \vec{f}(\vec{x}_0) + J(\vec{x}_0)(\vec{x}-\vec{x}_0)$$

← truncated vector-valued multivariable Taylor expansion.

~~The~~ Evaluating the above expression at $\vec{\xi}$ gives:

$$\vec{0} \approx \vec{f}(\vec{x}_0) + J(\vec{x}_0)(\vec{\xi}-\vec{x}_0)$$

$$\Rightarrow \quad \vec{\xi} \approx \vec{x}_0 - J^{-1}(\vec{x}_0)\,\vec{f}(\vec{x}_0)$$

And the resulting Newton iteration is: $\boxed{\vec{x}_{k+1} = \vec{x}_k - J^{-1}(\vec{x}_k)\,\vec{f}(\vec{x}_k)}$

It can be shown that convergence is also <u>quadratic</u>:

$$\|\vec{x}_{k+1} - \vec{s}\|_2 \approx \|\vec{x}_k - \vec{s}\|_2^2$$

What may go wrong? Much like Newton fails if $f'(s)=0$, we may have a ~~[illegible]~~ singular Jacobian $J(\vec{x}_k)$.

Other commonly seen form is:

$$J(\vec{x}_k)\left(\vec{x}_{k+1} - \vec{x}_k\right) = f(\vec{x}_k)$$

<u>or</u> solve $J(\vec{x}_k)\,\vec{s}_k = f(\vec{x}_k)$

set $\vec{x}_{k+1} = \vec{x}_k + \underbrace{\vec{s}_k}_{\text{step on iteration } k.}$

will revisit this when we talk about optimization,