Newton's Method in n-dimensions.

We have now seen that Newton's method (in 1-d) converges quadratically, at least near a solution.

Unlike the bisection method, Newton's method has a natural generalization to higher dimensions.

Given $f: \mathbb{R}^n \to \mathbb{R}^n$, we want to solve $f(\vec{x}) = \vec{0}$ for $\vec{x}$.

The derivative of f at $\vec{v}$ is the $n \times n$ Jacobian matrix

$$Df(\vec{x}) = \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & & \vdots \\ \partial f_n/\partial x_1 & \cdots & \partial f_n/\partial x_n \end{bmatrix}$$

In analogy to the Newton iteration

$$g(x) = x - \frac{f(x)}{f'(x)}$$

we construct

$$g(\vec{x}) = \vec{x} - \left(Df(\vec{x})\right)^{-1} f(\vec{x})$$

In practice, we write

$$g(\vec{x}) - \vec{x} = \vec{h}$$

and solve the linear system

$$Df(\vec{x})\,\vec{h} = -f(\vec{x})$$

to obtain an update step h.

<Demonstration for inverse Kinematics>

The analogue to the quadratic convergence theorem is

Newton-Kantorovich Theorem. (simplified)
Suppose $f: \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and for some K we have

$$\|Df(\vec{x}) - Df(\vec{y})\| \leq K \|\vec{x} - \vec{y}\|$$

for all $\vec{x}, \vec{y}$ in some convex set D_0.
Further, suppose we have some $\vec{x}_0 \in D_0$ so that $Df(\vec{x}_0)$ is invertible and has $\|Df^{-1}(\vec{x}_0)\| \leq \beta$, and $\|Df^{-1}(\vec{x}_0)\, f(\vec{x}_0)\| \leq \eta$.
Also, suppose

$$h = \beta K \eta \leq 1/2.$$

Further, define two numbers t_*, t_{**} by

$$t_* = \frac{1}{\beta K}\left(1 - \sqrt{1-2h}\right) \qquad t_{**} = \frac{1}{\beta K}\left(1 + \sqrt{1+2h}\right)$$

and suppose that the ball $B_{t_*}(\vec{x}_0)$ around $\vec{x}_0$ of radius t_* is contained in D_0.

Then: the Newton iteration

$$\vec{x}_{k+1} = \vec{x}_k - Df^{-1}(\vec{x}_k)\, f(\vec{x}_k)$$

defines a sequence of points which is well-defined, lies inside $B_{t_*}(\vec{x}_0)$, and converges to a solution x_* of $f(\vec{x}) = \vec{0}$.

Further, this solution is unique in the (larger) set $D_0 \cap B_{t_{**}}(\vec{x}_0)$, and if $h < 1/2$, the convergence is at least quadratic.

This is a lot to unpack! And even to parse, so we'll start by recalling

Some definitions, and facts.

~~Definition.~~ $\vec{x}-\vec{y}$, $f(\vec{x}_0)$ and $Df^{-1}(\vec{x}_0)f(\vec{x}_0)$ are vectors in $\mathbb{R}^n$, so we know what their norms are - the $\sqrt{\langle \vec{v},\vec{v}\rangle}$, as always.

$Df^{-1}(\vec{x})-Df^{-1}(\vec{y})$ and $Df^{-1}(\vec{x}_0)$ are nxn matrices. Their norms are operator norms, defined by

$$\|A\| = \sup\sqrt{\frac{\langle A\vec{x},A\vec{x}\rangle}{\langle \vec{x},\vec{x}\rangle}} = |\lambda_{max}|$$

the largest absolute eigenvalue of A.

Note that the eigenvalues of a symmetric nxn matrix A are real and that the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A. So the condition

$$\|Df^{-1}(\vec{x}_0)\| \leq \beta$$

~~and~~ can be rewritten

$$|\lambda_{max}(Df^{-1}(\vec{x}_0))| \leq \beta$$

or

$$\frac{1}{|\lambda_{min}(Df(x_0))|} \leq \beta$$

or

$$|\lambda_{min}(Df(x_0))| \geq \frac{1}{\beta}$$

Further,

$$\|Df^{-1}(\vec{x}_0)\, f(\vec{x}_0)\| \leq \|Df^{-1}(x_0)\|\, \|f(\vec{x}_0)\|.$$

~~So we can~~

$$\leq \frac{1}{|\lambda_{min}(Df(\vec{x}_0))|} \|f(\vec{x}_0)\|$$

$$\leq \beta \, \|f(\vec{x}_0)\|.$$

Therefore, we could weaken the theorem a little to ~~the~~

<u>Simplified NK Theorem</u>.

Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and on some convex $D_0 \subset \mathbb{R}^n$ there is some K so that

$$\| Df(\vec{x}) - Df(\vec{y}) \| \leq K \| \vec{x} - \vec{y} \|$$

Further, suppose that we have $\vec{x}_0 \in D_0$ ~~so that~~ and let λ_{min} be the smallest absolute eigenvalue of the matrix $Df(\vec{x}_0)$.

If $h = \frac{\|f(\vec{x}_0)\|}{\lambda_{min}^2} \cdot K \leq \frac{1}{2}$, ~~then~~ and

$B_{\frac{\lambda_{min}}{K}}(\vec{x}_0) \subset D_0$ then

$$\boxed{\begin{array}{l}\text{All Newton iterates} \\ \qquad X_{K+1} = X_K - Df(\vec{x}_K)^{-1} f(\vec{x}_K) \\ \text{are well defined, stay} \\ \text{in the ball } B_{\frac{\lambda_{min}}{K}}(\vec{x}_0), \\ \text{and converge at least} \\ \text{quadratically to some } \vec{x}_* \\ \text{with } f(\vec{x}_*) = \vec{0}.\end{array}}$$

Translation. If you have bounds on

a) how close $Df(\vec{x}_0)$ is to being singular

b) how fast $Df(\vec{x})$ can change as you move $\vec{x}$

c) how large $\|f(\vec{x}_0)\|$ is

then you can guarantee that Newton's method is going to work!

Note for the grad-school bound: The NK theorem is often used to prove the <u>existence</u> of a solution to $f(\vec{x})=\vec{0}$, and to <u>bound the location of the solution</u>, even when you don't care about computing it.

We now present a proof (due to J.M. Ortega)

Lemma 1. Let $\{\vec{y}_k\}$ be a sequence in $\mathbb{R}^n$ and $\{t_k\}$ be a sequence of nonnegative real numbers so that

$$\|\vec{y}_{k+1}-\vec{y}_k\| \leq t_{k+1}-t_k$$

and $t_k \to t_*$ with $t_* < \infty$. Then there is some $\vec{y}_* \in \mathbb{R}^n$ with $\vec{y}_k \to \vec{y}_*$ and

$$\| \vec{y}_k - \vec{y}_* \| \leq t^* - t.$$

Proof.

Notice that the sequence of t_k is increasing, since $t_{k+1} - t_k \geq \|\vec{y}_{k+1} - \vec{y}_k\| \geq 0$. Now

$$\begin{aligned}
\| \vec{y}_{k+p} - \vec{y}_k \| &\leq \sum_{i=1}^{p} \| \vec{y}_{k+i} - \vec{y}_{k+i-1} \| \\
&\leq \underbrace{\sum_{i=1}^{p} t_{k+i} - t_{k+i-1}}_{\text{telescoping sum}} \\
&\leq t_{k+p} - t_k \\
&\leq t_* - t_k
\end{aligned}$$

So the $\vec{y}_k$ are a Cauchy sequence and hence converge. Further,

$$\|\vec{y}_* - \vec{y}_K\| = \lim_{p \to \infty} \|\vec{y}_{K+p} - \vec{y}_K\| \leq t_* - t. \quad \square$$

To prove the next Lemma, we need to introduce some very cool ideas about matrices.

Recall that for $|r|<1$ we know

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$$

Making the substitution $x = 1-r$, we have the theorem: If $|1-x|<1$, then

$$\sum_{n=0}^{\infty} (1-x)^n = \frac{1}{x}$$

And if we substitute $x = PT$, then we get the result

"Banach Lemma": If $|1-PT|<1$ then

$$\sum (1-PT)^n P = \frac{P}{PT} = \frac{1}{T}.$$

Further,

$$\left|\frac{1}{T}\right| = \left|P \sum (1-PT)^n\right|$$

$$= |P| \left|\sum (1-PT)^n\right|$$

$$\leq |P| \sum |1-PT|^n$$

$$\leq \frac{|P|}{1-|1-PT|^n}.$$

Now here's the amazing thing! The entire argument works for matrices as well as numbers.

Banach Lemma. If T is an $n \times n$ matrix, T^{-1} exists if and only if there is an invertible $n \times n$ matrix P so that

$$\| I - PT \| < 1.$$

If T^{-1} exists,

$$T^{-1} = \sum_{n=0}^{\infty} (I - PT)^n P$$

and

$$\| T^{-1} \| \leq \frac{\| P \|}{1 - \| I - PT \|}.$$

This is really cool because it means you can <u>invert</u> matrices by summing powers of matrices.

We now apply the Banach Lemma to our N-K hypotheses.

Lemma. Assuming the hypotheses of the N-K theorem, for all $\vec{x}$ in $\mathbb{R}^n$ with $\|\vec{x}-\vec{x}_0\| < 1/\beta K$, $Df(\vec{x})$ is invertible and

$$\|Df^{-1}(\vec{x})\| \leq \frac{\beta}{1-\beta K\|\vec{x}-\vec{x}_0\|}$$

Proof. We will let $T = Df(\vec{x})$ and $P = Df^{-1}(\vec{x}_0)$. Then

$$\|I - PT\| = \|I - Df^{-1}(\vec{x}_0)\, Df(\vec{x})\|$$

$$= \|Df^{-1}(\vec{x}_0)\, Df(\vec{x}_0) - Df^{-1}(x_0)\, Df(\vec{x})\|$$

$$= \|Df^{-1}(x_0)\,[Df(\vec{x}_0) - Df(\vec{x})]\|$$

Now $\|AB\| \leq \|A\|\,\|B\|$ for any matrices (homework) so we have

$$\|Df^{-1}(\vec{x}_0)\left[Df(\vec{x}_0) - Df(\vec{x})\right]\| \leq \beta \|Df(\vec{x}_0) - Df(\vec{x})\|$$

$$\leq \beta K \|\vec{x}_0 - \vec{x}\| \leq \frac{\beta K}{\beta K} = 1.$$

The result now follows from the Banach lemma. $\square$

We now prove

Lemma. Assuming the hypotheses of the NK lemma, if we let

$$N(\vec{x}) := \vec{x} - Df^{-1}(\vec{x})\, f(\vec{x})$$

and $\vec{x}, N(\vec{x})$ ~~have~~ are within $1/\beta K$ of $\vec{x}_0$,

$$\|N(N(\vec{x})) - N(\vec{x})\| \leq \frac{1}{2}\, \frac{\beta K \|\vec{x} - N(\vec{x})\|^2}{1 - \beta K \|\vec{x}_0 - N(\vec{x})\|}$$

Proof. Notice that

$$f(\vec{x}) + Df(\vec{x})(N(\vec{x}) - \vec{x}) =$$

$$= f(\vec{x}) + Df(\vec{x})\left(\vec{x} - Df^{-1}(\vec{x})f(\vec{x}) - \vec{x}\right)$$

$$= f(\vec{x}) - f(\vec{x}) = \vec{0}$$

just by inserting the definition of N.

Making the substitution $\vec{x} \to N(\vec{x})$, we get

$$f(N(\vec{x})) + Df(N(\vec{x}))(N(N(\vec{x})) - N(\vec{x})) = \vec{0}$$

or

$$N(N(\vec{x})) - N(\vec{x}) = -Df^{-1}(N(\vec{x}))f(N(x))$$

Thus

$$\|N(N(\vec{x})) - N(\vec{x})\| = \|Df^{-1}(N(\vec{x}))f(N(\vec{x}))\|.$$

$$\leq \|Df^{-1}(N(\vec{x}))\| \, \|f(N(\vec{x}))\|$$

Now we already know

$$\|Df^{-1}(N(\vec{x}))\| \leq \frac{\beta}{1-\beta K\|\vec{x}_0 - N\vec{x}\|}$$

by the last lemma. And

$$\|f(N(\vec{x}))\| = \|f(N(x)) \underbrace{- f(\vec{x}) - Df(\vec{x})(N(\vec{x})-\vec{x})}_{\text{all this is zero!}}\|$$

At this point, we pause the proof for another awesome idea: the mean value theorem for vector-valued functions.