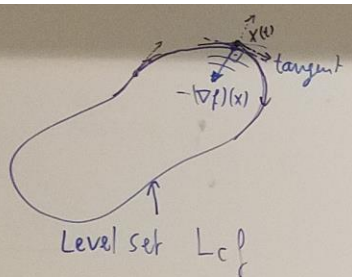
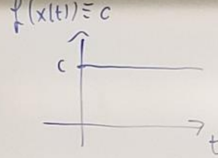
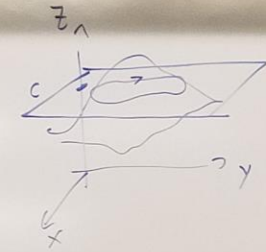






Consider a trajectory $t \mapsto x(t)$ contained in the level set $L_c f = \{x \in \mathbb{R}^n \mid f(x) = c\}$

$\Leftrightarrow \forall t \geq 0$ $(t \in [0,\infty))$ $f(x(t)) \equiv c = \text{const.} \Leftrightarrow \underbrace{\frac{d}{dt} f(x(t))}_{= (\nabla f)(x(t)) \cdot (\frac{d}{dt} x(t))} = 0$

Analysis of the rate / order of convergence of iterative optimization algorithms

Examples: Let $z_k = \left(\frac{1}{10}\right)^k$, then $z_1 = \frac{1}{10}$, $z_2 = \frac{1}{100}$, $z_3 = \frac{1}{1000}$, ... $z_k \xrightarrow{k\to\infty} \underbrace{0}_{z^*}$ or $\lim_{k\to\infty} z_k = \underbrace{0}_{z^*}$.

$$\lim_{k\to\infty} \frac{|z_{k+1} - z^*|}{|z_k - z^*|^1} = \lim_{k\to\infty} \frac{\left|\left(\frac{1}{10}\right)^{k+1} - 0\right|}{\left|\left(\frac{1}{10}\right)^k - 0\right|} = \frac{1}{10} = \delta$$

Linear convergence

$z_1 = 0.1$, $z_2 = 0.01$, $z_3 = 0.001$, ... $z^* = 0.000000\ldots$

$\Rightarrow$ Since $\delta = 0.1$, every iteration adds another digit of accuracy to the objective value

If $\delta = 0.9$, every 22 iterations to add another digit of accuracy, because $(0.9)^{22} \approx 0.1$

The smaller the δ, the better the convergence

Let $z_k = \left(\frac{1}{10}\right)^{2^k}$, then $z_1 = 0.01$, $z_2 = 0.0001$, $z_3 = 0.00000001$, ...

$$\lim_{k\to\infty} \frac{|z_{k+1} - z^*|}{|z_k - z^*|^2} = \lim_{k\to\infty} \frac{\left|\left(\frac{1}{10}\right)^{2^{k+1}} - 0\right|}{\left|\left(\frac{1}{10}\right)^{2^k} - 0\right|^2} = \frac{\left(\frac{1}{10}\right)^{2^k} \left(\frac{1}{10}\right)^{2^k}}{\left(\frac{1}{10}\right)^{2^k} \cdot \left(\frac{1}{10}\right)^{2^k}} = 1$$

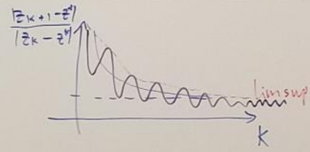
Quadratic convergence

Let $z_k = \frac{1}{k!}$, $z_k \xrightarrow{k\to\infty} z^* = 0$

$$\lim_{k\to\infty} \frac{|z_{k+1} - z^*|}{|z_k - z^*|^1} = \lim_{k\to\infty} \frac{\left|\frac{1}{(k+1)!} - 0\right|}{\left|\frac{1}{k!} - 0\right|} = \lim_{k\to\infty} \underbrace{\frac{1}{k+1}}_{\delta_k} = 0$$

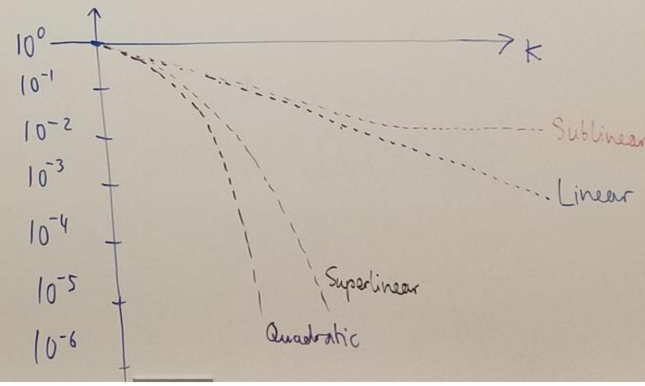
Superlinear convergence

Definition (Order of Convergence)



Let $\beta = \limsup_{k\to\infty} \dfrac{|z_{k+1} - z^*|}{|z_k - z^*|^p}$.

1. Linear convergence: $p=1$ and $0<\beta<1$.
2. Superlinear convergence: $p=1$, $\beta = 0$
3. Sublinear convergence: $p=1$, $\beta=1$
4. Quadratic convergence: $p=2$, $\beta<\infty$



Theorem Linear convergence of steepest descent

Let $f(x) = \frac{1}{2}x^T Q x - c^T x$, $Q \succ 0$

and let M and m denote the largest and smallest eigenvalues of Q, respectively.

Then
$$\frac{f(x^{k+1}) - f(x^*)}{(f(x^k) - f(x^*))^1} \leq \left(\frac{\frac{M}{m} - 1}{\frac{M}{m}+1}\right)^2.$$

$K(Q)$	upper bound on δ	# iter. to reduce gap by 0.1
1.1	0.0023	1
3.0	0.25	2
10.0	0.67	6
100.0	0.96	58
		⋮

Remarks:
- $K(Q) = \frac{M}{m} \geq 1$ is the condition number of Q.
- $K(Q)$ plays an important role in analyzing computations involving Q.
- The number of iterations needed to reduce the optimality gap by an order of magnitude grows linearly by $K(Q)$.

Theorem Quadratic convergence of Newton's Method

Let $f \in C^3$ on $\mathbb{R}^n$ and assume that at the local minimum x^* we have $(Hf)(x^*) \succ 0$. Then if initialized sufficiently close to x^*, the points generated by Newton's Method converge quadratically to x^*.