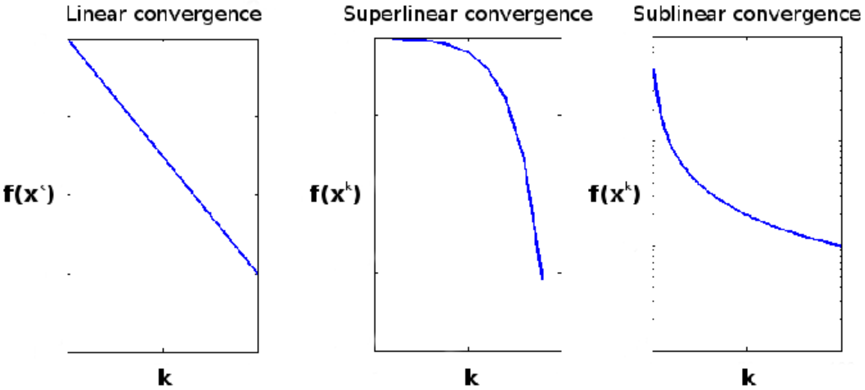
Convergence Analysis

Instructor: Prof. Ganesh Ramakrishnan

Convergence



R-convergence

- Let us consider the convergence result we got by assuming Lipschitz continuity with backtracking and exact line searches:

(& convexity)

$$f(x^k) - f(x^*) \leq \frac{\left\| x^{(0)} - x^* \right\|^2}{2tk}$$

- We will characterize this using **R-convergence**
- 'R' here stands for 'root', as we are looking at convergence rooted at x^*

Q-convergence

- We say that the sequence $s^1, \dots, s^k$ is **R-linearly** convergent if $\left\|s^k - s^*\right\| \leq v^k$, $\forall k$, and $\left\{v^k\right\}$ converges **Q-linearly** to zero
- $v^1, \dots, v^k$ is Q-linearly convergent if

$$\frac{\left\|v^{k+1} - v^*\right\|}{\left\|v^k - v^*\right\|} \leq r \in (0,1)$$

for some $k \geq \theta$, and $r \in (0,1)$

  - 'Q' here stands for 'quotient' of the norms as shown above

R-convergence assuming Lipschitz continuity

- Consider $v^k = \frac{\left\|x^{(0)} - x^*\right\|^2}{2tk} = \frac{\alpha}{k}$, where α is a constant
- Here, we have $\frac{\left\|v^{k+1} - v^*\right\|}{\left\|v^k - v^*\right\|} \leq \frac{K}{K+1}$, where K is the final number of iterations
  - $\frac{K}{K+1} < 1$, but we don't have $\frac{K}{K+1} < r$
- Thus, $v^k = \frac{\alpha}{k}$ is not Q-linearly convergent as there exist no $v < 1$ s.t.
$\frac{\alpha/(k+1)}{\alpha/k} = \frac{k}{k+1} \leq v, \forall k \geq \theta$
- Strictly speaking, for Lipschitz continuity alone, gradient descent is not guaranteed to give R-linear convergence
- In practice, Lipschitz continuity gives “almost” R-linear convergence – not too bad!

R-convergence assuming Strong convexity

- Now, let us consider the convergence result we got by assuming Strong convexity with backtracking and exact line searches:

$$f(x^k) - f(x^*) \leq \left(1 - \frac{m}{M}\right)^k \left(f(x^{(0)}) - f(x^*)\right)$$

Boyd uses M instead of L

- Here, v^k can be considered $\left(1 - \frac{m}{M}\right)^k \alpha$
  - $v^* = 0$
- We get

$$\frac{v^{k+1} - v^*}{v^k - v^*} = \left(1 - \frac{m}{M}\right) \in (0, 1)$$

  - We now have an upper bound < 1, unlike before
- As $r = \left(1 - \frac{m}{M}\right) \in (0, 1)$, v^k is Q-linearly convergent
  - Thus, under strong convexity, gradient descent is R-linearly convergent

- *Question:* Is gradient descent under Strong convexity also Q-linearly convergent?
- Recall one of the intermediate steps in getting the convergence results:
  $f(x^{k+1}) - f(x^*) \leq \left(1 - \frac{m}{M}\right)\left(f(x^k) - f(x^*)\right)$
  - $\implies \frac{f(x^{k+1}) - f(x^*)}{f(x^k) - f(x^*)} \leq \left(1 - \frac{m}{M}\right)$
- Now, $r = \left(1 - \frac{m}{M}\right) \in (0, 1)$
- Yes, gradient descent under Strong convexity is also Q-linearly convergent

- Taking hint from this analysis, if Q-linear,

$$\frac{\left\|s^{k+1} - s^*\right\|}{\left\|s^k - s^*\right\|} \leq r \in (0,1)$$

then,
$\left\|s^{k+1} - s^*\right\| \leq r\left\|s^k - s^*\right\|$
$\leq r^2\left\|s^{k-1} - s^*\right\|$
$\vdots$
$\leq r^k\left\|s^{(0)} - s^*\right\|$, which is v^k for R-linear

- Thus, Q-linear convergence $\implies$ R-linear convergence
  - Q-linear is a special case of R-linear
  - R-linear gives a more general way of characterizing linear convergence
- Q-linear is an 'order of convergence' → Linear | Quadratic

  r is the 'rate of convergence' → For Lipschitz + strong, r is small if $m/L \to 1$

- Q-superlinear convergence:

$$\lim_{k\to\infty} \frac{\left\|s^{k+1} - s^*\right\|}{\left\|s^k - s^*\right\|} = 0$$

- Q-sublinear convergence:

$$\lim_{k\to\infty} \frac{\left\|s^{k+1} - s^*\right\|}{\left\|s^k - s^*\right\|} = 1$$

  - *e.g.* For Lipschitz continuity, v^k in gradient descent is Q-sublinear: $\lim_{k\to\infty} \frac{k}{k+1} = 1$

- Q-convergence of order p: Q: Does $p \geq 2 \Rightarrow$ superlinear?

Order of convergence

$$\forall k \geq \theta, \frac{\left\|s^{k+1} - s^*\right\|}{\left\|s^k - s^*\right\|^p} \leq M$$

  - *e.g.* $p = 2$ for Q-quadratic, $p = 3$ for Q-cubic, *etc.*

- **Claim:** Q-convergences of the order p are special cases of Q-superlinear convergence
- $\forall k \geq \theta$, $\frac{\left\| s^{k+1} - s^* \right\|}{\left\| s^k - s^* \right\|^p} \leq M$

$$\implies \lim_{k \to \infty} \frac{\left\| s^{k+1} - s^* \right\|}{\left\| s^k - s^* \right\|} \leq \lim_{k \to \infty} M \left\| s^k - s^* \right\|^{p-1} = 0$$

- Therefore, irrespective of the value of M (as long as $M \geq 0$), order $p > 1$ implies Q-superlinear convergence