Line Search → Analysis

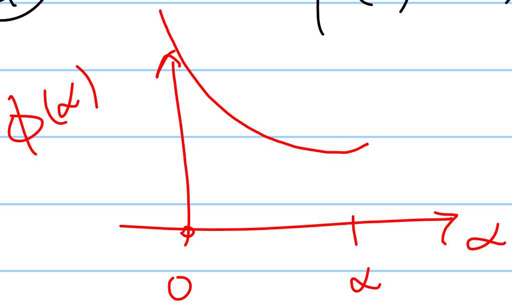
Recap →

① Backtracking line search

Repeat until $\left[f(x_k + \alpha p_k) \le f(x_k) + c_1 \alpha \nabla f_k^T p_k \right]$ at $\alpha = 0$

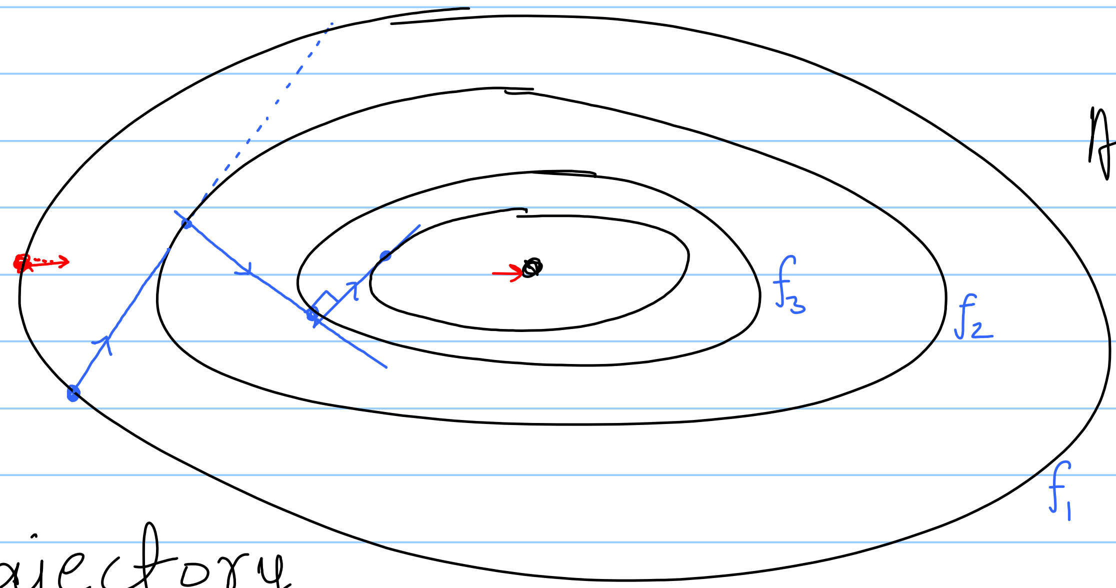
$\alpha \leftarrow \alpha \rho$, $\rho \in (0,1)$

② $\phi(\alpha) \rightarrow \phi'(\alpha) \leftarrow$ calculate pick any α



$\left.\begin{array}{l} \phi(0) \\ \phi'(0) \end{array}\right\}$ fit a quadratic & minimize

Observe?



Assume $f_1 > f_2 > f_3$

↳ $p_k \to -\nabla f_k$ (grad descent)

↳ Exact line searches

① Zig-zag trajectory

② step sizes are decreasing

③ Each step is $\perp^r$ to previous step!

$$s_k = x_{k+1} - x_k \quad ①$$

#3 $\rightarrow$ $s_k^T s_{k+1} = 0$

$$\boxed{s_{k+2}^T s_{k+1} = 0}$$

①
$$x_{k+1} = x_k + \alpha_k p_k$$

$$\Rightarrow s_k = \alpha_k p_k$$

$$\phi(\alpha) = f(x_k + \alpha p_k)$$

Exact line search

$$\frac{d\phi}{d\alpha} = \nabla f(x_k + \alpha p_k)^T p_k$$

$$\Rightarrow \frac{d\phi}{d\alpha} = 0 \Rightarrow \nabla f(x_k + \alpha_k p_k)^T p_k = 0 \quad ②$$

③
$$p_{k+1} = -\nabla f_{k+1} = -\nabla f(x_k + \alpha_k p_k)$$

$$s_{k+2}^T s_{k+1} = \alpha_{k+2} \alpha_{k+1} p_{k+1}^T p_k$$

$$= -\alpha_{k+2} \alpha_{k+1} \nabla f(x_k + \alpha_k p_k)^T p_k = 0$$

Convergence & Rate

Yes! Wherever I start → I reach a stationary pt.

a) Define θ_k → $\cos\theta_k = \dfrac{-\nabla f_k^T p_k}{\|\nabla f_k\| \, \|p_k\|}$

$\theta_k = 0 \longrightarrow$ Steepest descent.

$-90 < \theta_k < 90$ IN General

$\cos\theta_k > 0$

(b) Zoutendijk Condition.

$$x_{k+1} = x_k + \alpha_k p_k$$

α_k: Satisfies Wolfe Conditions

p_k: descent dir (1)

Conditions of the thm:

- fn should be bounded from below.
- f is continuously differentiable
- the gradient of f is Lipschitz conts

$$\Rightarrow \|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \forall x, y \in N$$

Then:

$$\underbrace{\sum_{j=0}^{k} \cos^2(\theta_j) \|\nabla f_j\|^2 < \infty}$$

Curvature condition

$$-\nabla f_{k+1}^T p_k \leq -c_2 \nabla f_k^T p_k$$

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \underbrace{\nabla f(x_k)^T p_k}$$ Sub $\nabla f_k^T p_k$

$$\left[(\nabla f_{k+1} - \nabla f_k)^T p_k\right] \geq (c_2 - 1) \nabla f_k^T p_k. \quad — ①$$

Talk about Lipschitz. $\left[(\nabla f_{k+1} - \nabla f_k)^T p_k\right] \leq \|\nabla f_{k+1} - \nabla f_k\| \, \|p_k\|$

$$\leq L \, \|x_{k+1} - x_k\| \, \|p_k\|$$

$$= L \alpha_k \|p_k\|^2 \quad — ②$$

Combine ① & ②

$$L\alpha_k \|p_k\|^2 \geq (c_2 - 1) \nabla f_k^T p_k$$

$$\alpha_k \geq \left[\frac{c_2 - 1}{L} \frac{\nabla f_k^T p_k}{\|p_k\|^2} \right] \quad \text{--- } (3)$$

Sufficient decrease

$$\longrightarrow \quad f_{k+1} \leq f_k + c_1 \alpha_k \nabla f_k^T p_k$$

$$\leq f_k + \frac{c_1 (c_2 - 1)}{L} \frac{(\nabla f_k^T p_k)^2}{\|p_k\|^2}$$

$$\leq f_k - c \underbrace{\frac{(\nabla f_k^T p_k)^2}{\|p_k\|^2}}$$

$$c = \frac{c_1 (1 - c_2)}{L}$$

$$f_{k+1} \leq f_k - c \cos^2\theta_k \|\nabla f_k\|^2$$

Telescoping.

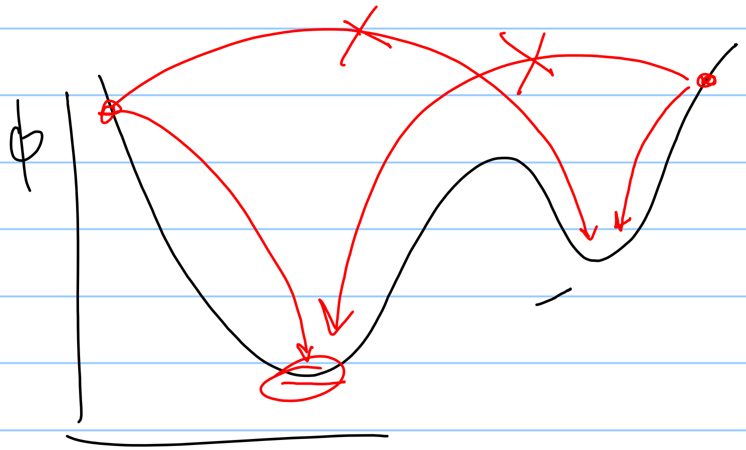
$$f_{k+1} \leq f_0 - c \sum_{j=0}^{k} \cos^2\theta_j \, \|\nabla f_j\|^2$$

$\because$ f is bounded from below

$$\sum_{j=0}^{k} \cos^2\theta_j \, \|\nabla f_j\|^2 < \infty$$

QED

$\Downarrow$



Corollary $\rightarrow$ $\lim_{k\to\infty} \cos^2\theta_k \, \|\nabla f_k\|^2 \to 0$

p_k is a descent dir $\Rightarrow$ $\cos\theta_k \neq 0$

$\Rightarrow \lim_{k\to\infty} \|\nabla f_k\| \to 0$

i.e. a stationary point.