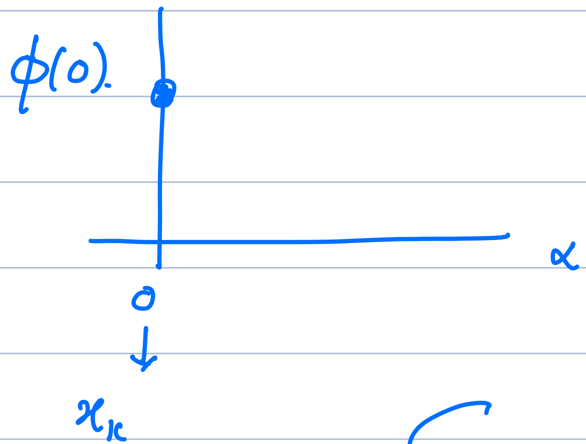
Backtracking line search

repeat until $\left[f(x_k + \alpha p_k) \le f(x_k) + c\alpha \nabla f_k^T p_k \right]$

$\alpha \leftarrow \alpha \rho$

where $0 < \rho < 1$. Typically $\alpha^{(0)} = 1$.

$$\phi(\alpha) = f(x_k + \alpha p_k)$$
$$\phi'(\alpha) = \nabla f(x_k + \alpha p_k)^T p_k$$
$$\phi'(0) = \nabla f(x_k)^T p_k$$



3 pieces of data

1) $\phi(0)$, 2) $\phi'(0)$, 3) $\phi(\alpha^{(0)})$

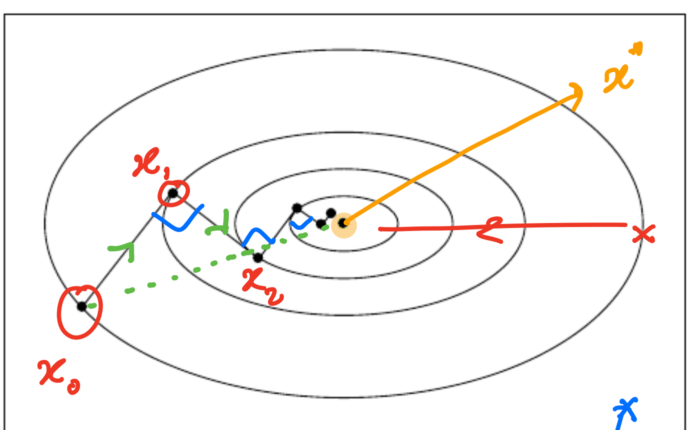
↳ can fit a quadratic fn: $q(\alpha)$

↳ Then? find the minima.

Interpolation to find a poly approx of $\phi(\alpha)$.

Line Search - Analysis

(1)



1) zigzag

2) stepsizes are decreasing.

3) angles are 90° betn steps.

↳ for exact line search.

$$s_{k+1} = x_{k+1} - x_k \quad \therefore \text{ obsr \#3} \Rightarrow$$

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

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

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

$\phi(\alpha) = f(x_k + \alpha_k p_k)$. If line search exact:

$$\Rightarrow \boxed{\phi'(\alpha) = 0} = \nabla f(x_k + \alpha_k p_k)^T p_k = \nabla f(x_{k+1})^T p_k = 0$$

$$\left.\begin{array}{l} p_k = -\nabla f_k \\ p_{k+1} = -\nabla f_{k+1} \end{array}\right\}$$

$$\Rightarrow s_{k+2}^T s_{k+1} = \alpha_k \alpha_{k+1} \left(p_{k+1}^T p_k \right) = 0.$$

② Convergence & Rate.

↳ Yes! → It achieves global convergence.

⇒ wherever I start → I reach a stationary point.

a) $$\cos \theta_k = \frac{(-\nabla f_k^T) p_k}{\|\nabla f_k\| \, \|p_k\|}$$ $\theta_k = 0 \rightarrow$ gradient descent.

If descent direction $-90 < \theta_k < 90$

$$\cos\theta_k \geq 0 .$$

b) Zoutendijk - Thm. $x_{k+1} = x_k + \alpha_k p_k$

descent dir

satisfies wolfe cond^n s

↳ f is bounded from below

↳ f is continuously differentiable.

↳ the gradient of f is Lipschitz cont^n s.

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

Curvature cond^n s :

$$\nabla f(x_k + \alpha p_k)^T p_k \geq c_2 \nabla f(x_k)^T p_k$$

Subtr $\nabla f^T p_k$ both sides.

$$\underbrace{\left[\nabla f(x_k + \alpha p_k)^T - \nabla f(x_k)\right] p_k} \geq (c_2 - 1) \nabla f(x_k)^T p_k . \quad (1)$$

$$(\nabla f_{k+1} - \nabla f_k)^T p_k \leq \|\nabla f_{k+1} - \nabla f_k\| \, \|p_k\|$$

$$\leq L \alpha_k \|p_k\|^2 \quad (2)$$

$$\Rightarrow \boxed{\alpha_k \geq \frac{c_2 - 1}{L} \frac{\nabla f_k^T p_k}{\|p_k\|^2}} \quad (3)$$

↳ Sufficient decrease:

$$f_{k+1} \leq f_k + c_1 \alpha \nabla f_k^T p_k \quad — \quad (4)$$

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

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

Sum both sides 0 to k

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

$$c \sum_{j=0}^{k} \cos^2\theta_j \|\nabla f_j\|^2 \leq \underbrace{f_0 - f_{k+1} \sim M}_{\text{+ve quantity.}}$$

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

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

$f_1 < f_0$
$f_2 < f_1$

Ensure p_k is a legit descent dir

$$\Rightarrow \quad |\cos\theta_k| > \delta$$

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

$\Rightarrow$ a stationary point.