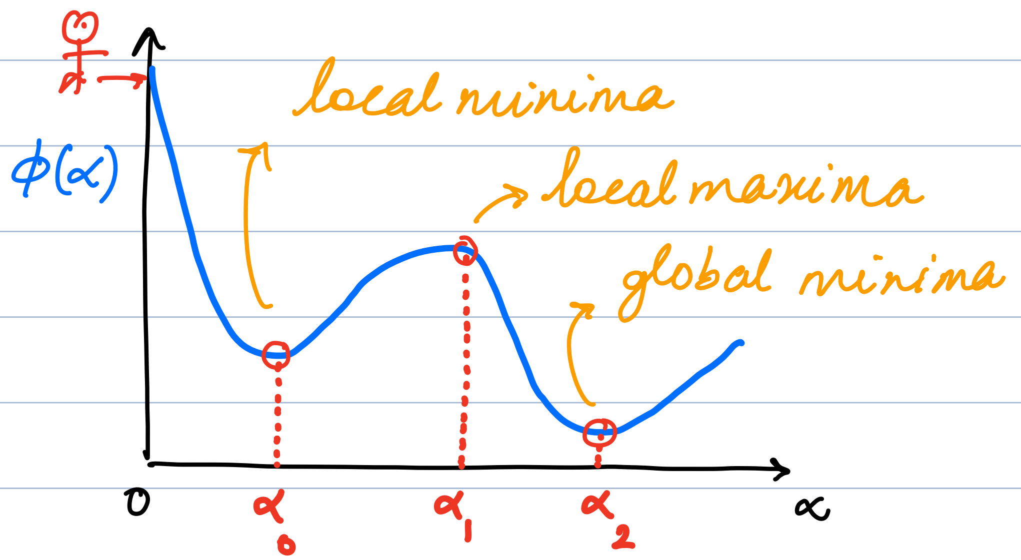
Line Search - Wolfe Conditions

→ $x_{k+1} = x_k + \underbrace{\alpha}_{\text{TBD}} \underbrace{p_k}_{\text{fixed}}$

$\because$ p_k is a descent dir, $\alpha > 0$

↪ $p_k^T \nabla f_k < 0$.

$\phi(\alpha) = f(x_k + \alpha p_k)$, $\alpha > 0$



$\phi'(\alpha) = 0$

→ We do an inexact line search

⇓

Wolfe conditions.

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

→ aside: $f(x, y)$.

$x = x_0 + t p$

$y = y_0 + t q$

$$f(x(t), y(t))$$

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

$$= \frac{\partial f}{\partial x}\, p + \frac{\partial f}{\partial y}\, q$$

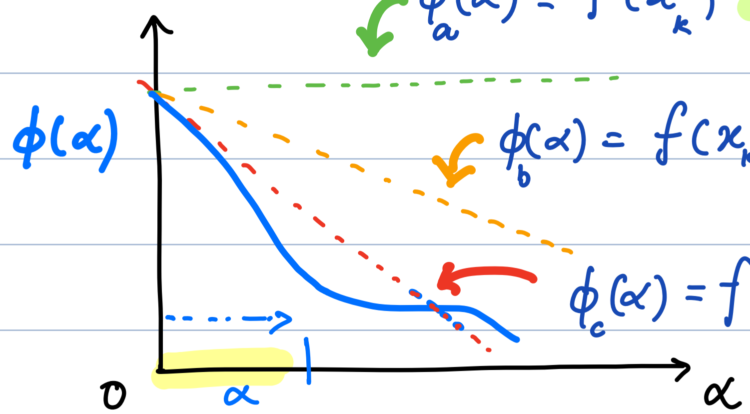
$$= (\nabla f)^T \begin{pmatrix} p \\ q \end{pmatrix}$$

① "Sufficient Decrease"

I get info of $\phi(0)$ and $\phi'(0)$

→ We can construct 1^{st} order Taylor series.

i.e. $\quad l(\alpha) = f(x_k) + \alpha \underline{\nabla f_k^T p_k}$.



$\phi_a(\alpha) = f(x_k)$

$c_1 = 0$

$\phi_b(\alpha) = f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k$

$c_1 = 1$

$\phi_c(\alpha) = f(x_k) + \alpha \nabla f(x_k)^T p_k$

$$f(x_k + \alpha p_k) \leq f(x_k) + c_1 \alpha \nabla f(x_k)^T p_k$$

$$c_1 \in (0,1)$$

"<u>Armijo rule</u>"

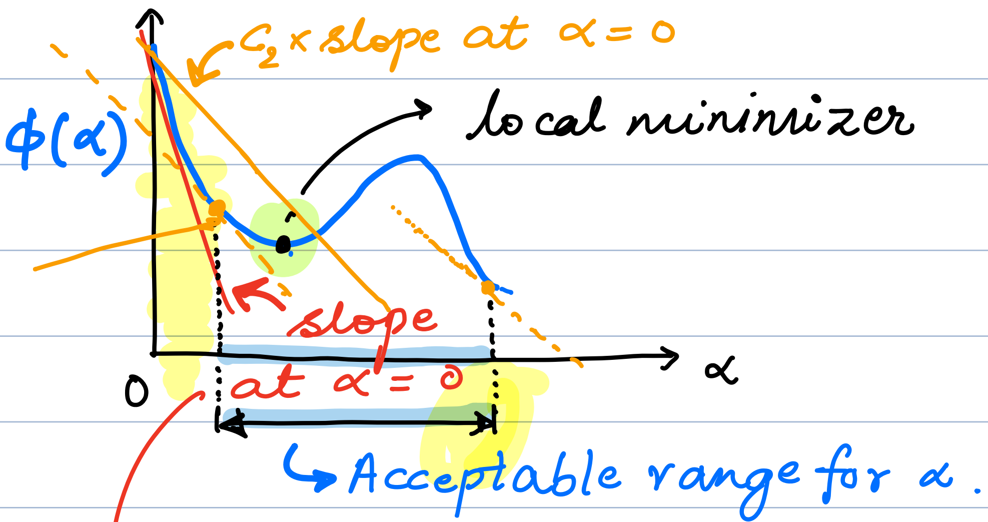
Problem → It allows very small step sizes. $\sim 10^{-4}$

② "Curvature Condition"

$\propto \phi'(\alpha)$

Goal: $\nabla f(x_k + \alpha p_k) = 0$

$$|\phi'(\alpha)| < \text{slope of linear approx at } x_k.$$



$(\nabla f_k^T p_k)$

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

$$\underbrace{-\nabla f(x_k + \alpha p_k)^T p_k}_{<0 \text{ for small } \alpha.} \leq -c_2 \boxed{\nabla f(x_k)^T p_k} \rightarrow < 0$$

$0 < c_1 < c_2 < 1$

Slope here at α < slope at 0

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

$(0.1 - 0.9)$

$$\leq -\left(\tfrac{1}{2}\right)(-100)$$

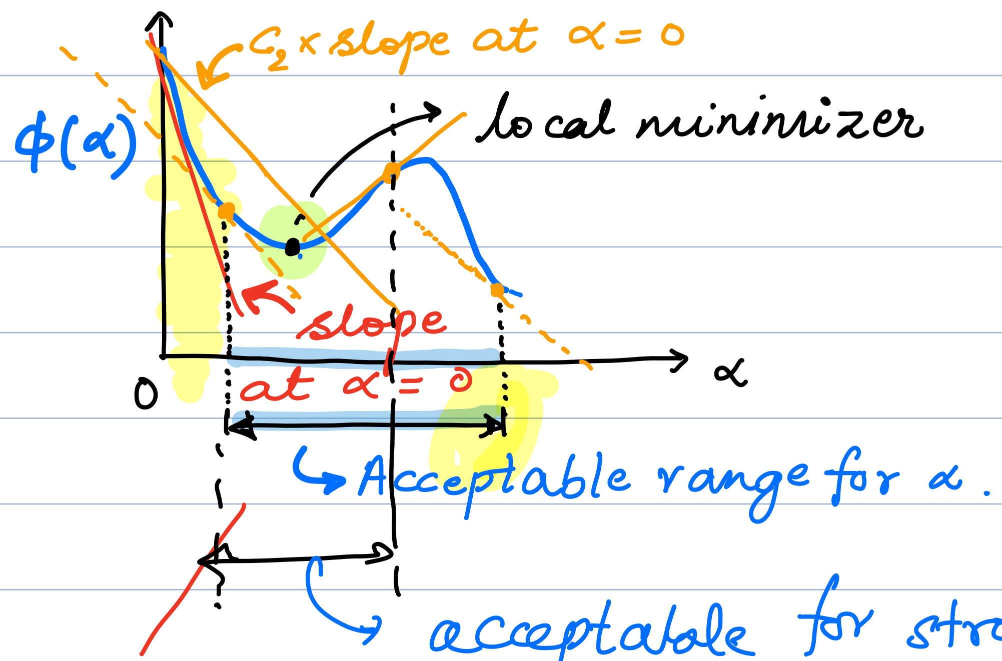
$$0 \leq 50$$

↳ Once we go past a stationary point then $\phi'(\alpha) > 0$

$$\underbrace{-\underbrace{\phi'(\alpha)}_{>0}}_{<0} \leq \underbrace{-c_2 \underbrace{\phi'(0)}_{<0}}_{>0}$$

The fix : Strong Wolfe condn.

$$|\phi'(\alpha)| \leq c_2 |\phi'(0)| \quad , \quad c_1 < c_2 < 1$$



↳ acceptable for strong Wolfe condn

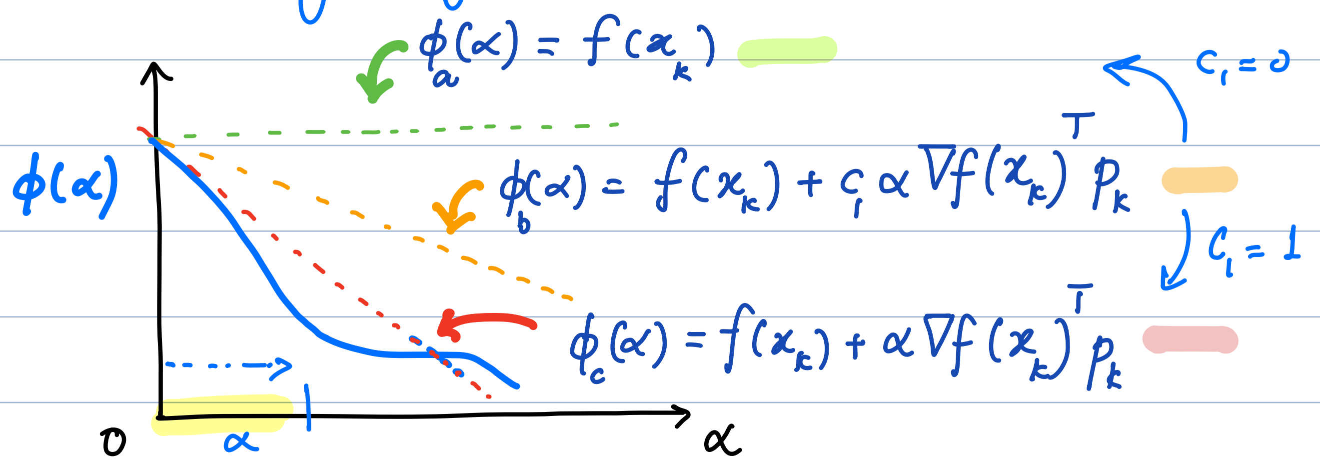
Wolfe condns.

1) Sufficient decrease
2) Curvature condn (regular or strong).

↳ Goldstein condns

→ Backtracking line search.

Use only sufficient decrease + "backtrack"



→ Start with a large α

Repeat until $f(x_k + \alpha p_k) \leq f(x_k) + c\alpha \nabla f_k^T p$

$\alpha \leftarrow \alpha \times \rho$

where $0 < \rho < 1$