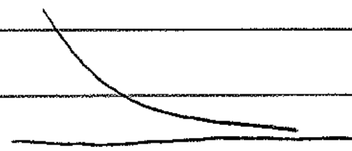
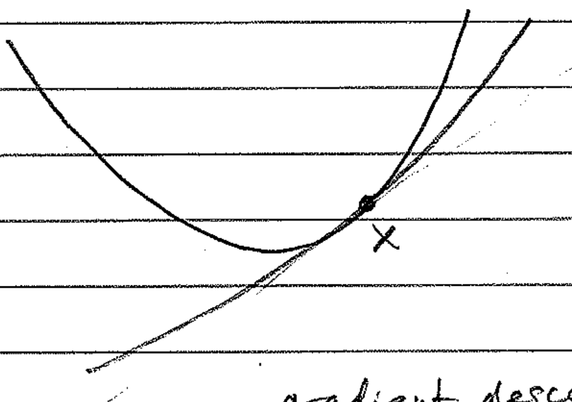
$f''(x) > 0$ strict convexity

$f''(x) \geq m$

x^4

e^{-x}





$\hat{f}(y)$

$$f(y) \approx \boxed{f(x) + \nabla f(x)^T (y-x) + \frac{1}{2t} \|y-x\|_2^2}$$

usual: $\frac{1}{2}(y-x)^T \nabla f^2(x) (y-x)$

gradient descent

chooses update

$$x^+ = \underset{y}{\operatorname{argmin}} \ \hat{f}(y)$$

$$= \operatorname{argmin} \ f(x) + \nabla f(x)^T (y-x) + \frac{1}{2t} \|y-x\|_2^2$$

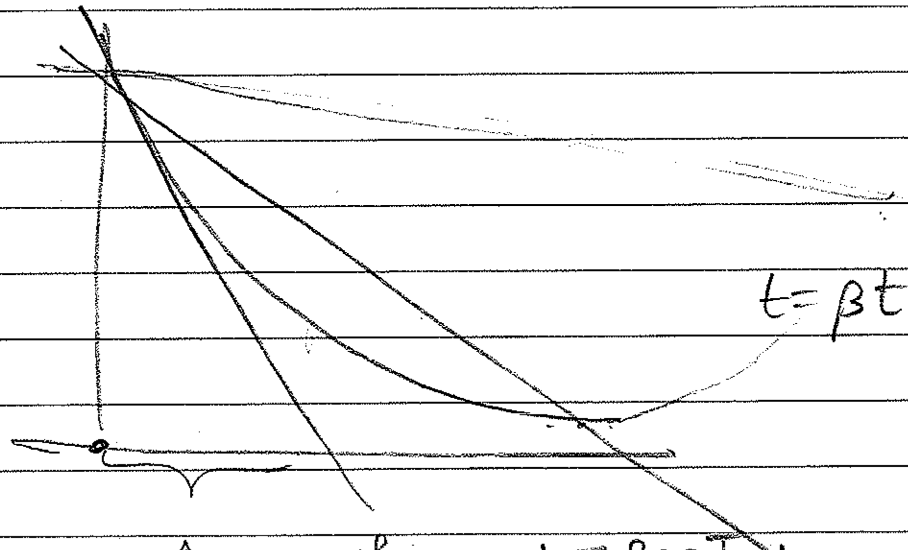
$$= \frac{1}{2t} y^T y + (\nabla f(x) - \tfrac{1}{t} x)^T y + \text{cons.}$$

$f(u) = au^2 + bu + c$, $-b/2a$

$f(u) = u^T A u + b^T u + c$, $-\frac{1}{2} A^{-1} b$

minimizer

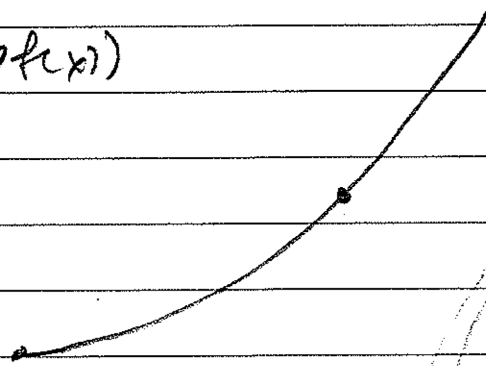
$$\underline{x^+ = x - t \nabla f(x)}$$

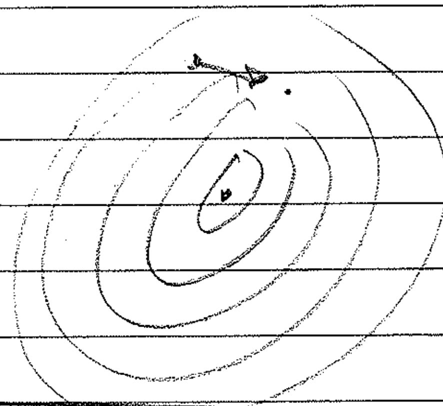


Δ $\qquad f(x) + t\nabla f(x)^T \Delta$

$f(x) + \alpha t \nabla f(x)^T \Delta$

$g(s) = f(x - s\nabla f(x))$





---

f to within ε

$f(x^{(k)}) - f(x^*) \le \frac{C}{k} = \varepsilon \qquad O(1/k)$

$k = \frac{C}{\varepsilon} \qquad O(1/\varepsilon)$