min $f(x)$.

where $f(x)$ convex. 2^{nd} continuous differentiable

- Produce sequence of points $x^{(k)} \in \mathrm{dom}\, f$

  with $f(x^{(k)}) \to p^*$ where $p^* = \inf_x f(x)$

1. Initial point & Sublevel set

Starting Point $\begin{cases} x^{(0)} \in \mathrm{dom}\, f \\ \text{sublevel set } S = \{x \mid f(x) \le f(x^{(0)})\} \text{ is closed} \end{cases}$

Strong convexity $\iff \underline{\nabla^2 f(x) \succeq mI}$. $\exists m > 0$

- Implication

$$f(y) \ge f(x) + \nabla f(x)^T (y - x) + \frac{m}{2}\|x - y\|_2^2$$

we minimize rhs for y

$$\Rightarrow f(x) - p^* \le \frac{1}{2m}\|\nabla f(x)\|_2^2$$ useful as stopping criterion

2. Descent Methods

Search Direction

$$x^{(k+1)} = x^{(k)} + \underbrace{t^{(k)}}_{\text{Step size}} \Delta x^{(k)} \quad \text{with } f(x^{(k+1)}) < f(x^{(k)})$$

- From convexity. we know $\underline{\nabla f(x)^T \Delta x < 0}$ Descent Direction

① Linear Search Type

$$t = \operatorname*{argmin}_{t > 0} f(x + t\Delta x)$$ ~ Exact Line Search

$\Rightarrow$ Backtracking line search

$t = 1$ at start. repeat $t := \beta t$ until

$$f(x + t\Delta x) < f(x) + \underline{\alpha t \nabla f(x)^T \Delta x} \quad \alpha \in (0, \tfrac{1}{2})$$

Degrade the slope

② Gradient Descent Method

$\Delta x = -\nabla f(x)$

- Stopping Criterion $\|\nabla f(x)\|_2 \leq \varepsilon$

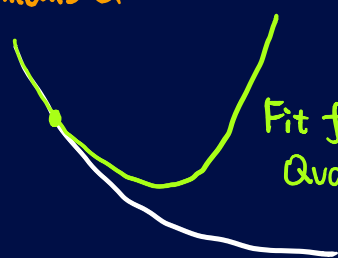
③ Steepest Descent Method

* Here we Normalized

$\Delta x_{nsd} = \operatorname{argmin} \{ \nabla f(x)^T v \mid \|v\| = 1 \}$ ~ Most steep downhill

or $\Delta x_{sd} = \|\nabla f(x)\|_* \, \Delta x_{nsd}$ *Unnormalized

3. Newton Step



Fit $f(x)$ with Quadratic!

$\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x)$

- $x + \Delta x_{nt}$ minimizes second-order approximation

$$\hat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v$$

- Newton Decrement

$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x) \right)^{\frac{1}{2}}$$

gives a measure of proximity of x to x^*

Newton's method

1. Compute $\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x)$, $\lambda^2 = \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)$
2. quit if $\frac{\lambda^2}{2} \leq \varepsilon$
3. choose step size t by $x := x + t\Delta x_{nt}$