Interior Point Method I

Tuesday, January 31, 2023 11:17 PM

Given: $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ $\quad L = \langle c, A, b \rangle$

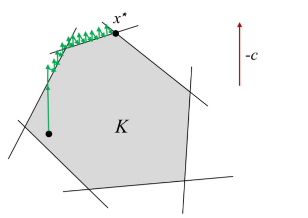
- Linear Programming

Find $x^* \in \underset{x}{\operatorname{argmin}}: c^T x$ s.t. $Ax \geq b$ $\qquad K = \{x \mid Ax \geq b\}$

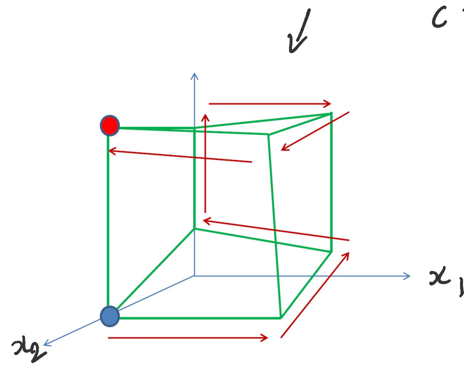
- Gradient Descent / Simplex Method

$f(x) = c^T x$

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



(a) Gradient Descent optimization path.



$c = (0, -1)$

$c^T x = -x_2$

- Reduction to Unconstrained Optimization

$$\min: c^T x \quad Ax \geq b \qquad \min_x : w \cdot c^T x + B(x).$$

$$\begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \quad 1 \leq i \leq m \quad a_i^T x - b_i \geq 0 \qquad B(x) = \sum_{i=1}^{m} -\log(a_i^T x - b_i)$$

$$\boxed{F_w(x) = w \cdot (c^T \cdot x) + B(x)}$$

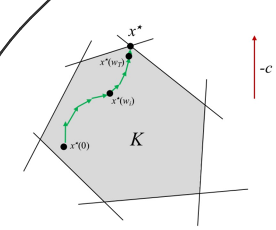
$$x^*(w) = \underset{x}{\operatorname{argmin}}\ F_w(x).$$

Goal: start at $w = 0$ & keep increasing w until very near to x^*.

- Interior Point Method

$w = 0$

1. Start at (or really close to) $x^\star(0)$.
2. Choose a sequence of geometrically increasing values $w_1, \ldots, w_T$ with $w_{i+1} = (1+q)w_i$.
3. For $i = 1, \ldots, T$, use **Newton**'s method to find $x^\star(w_i) = \arg\min F(w_i)$, using $x^\star(w_{i-1})$ as a starting point.[2]
4. Return $x^\star(w_T)$.



(b) Ideal Interior Point optimization path.

Q1: $x^*(0) = \underset{x}{\operatorname{argmin}}\ F_0(x) = \underset{x}{\operatorname{argmin}}\ B(x)$

Find a feasible point $\tilde{t} = \arg\min: t$

$Ax \geq b(1-t)$

$0 \leq t \leq 1$

$x = \bar{0}$, $t = 1$ is feasible.

if $\tilde{t} = 0$ then found x s.t. $Ax \geq b$

else $Ax \geq b$ is infeasible.

found x s.t. $Ax \geq b$ $\qquad 0 \leq t \leq$

Else $Ax \geq b$ is infeasible.

Q2: What should be w_1, w_T, q?

$w_1 = 2^{-O(L)}$ suffices. $\qquad \boxed{w_T = \frac{n}{\varepsilon}}$ then $c^T \cdot x^*(w_T) \leq c^T x^* + \varepsilon$

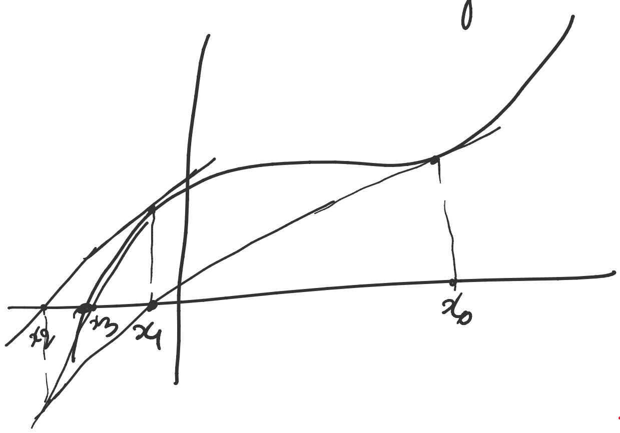
$$T \leq O\left(\frac{L}{q} \lg \frac{n}{\varepsilon}\right)$$

GD to get ε- approx for convex min $\lg\left(\frac{1}{\varepsilon}\right)$

Detour: Newton's Method

- Single Variate (root finding) $\qquad g: \mathbb{R} \to \mathbb{R}$.

$$g'(x) = \lim_{y \to x} \frac{g(y) - g(x)}{(y - x)}$$



$$0 = g(y) = g(x) + g'(x)(y - x)$$

$$y = x - \frac{g(x)}{g'(x)}$$

Thm: (Quadratic Convergence)

Let g be twice differentiable & g'' is continuous.

$r \in \mathbb{R}$ be the root & x_0 start point.

$$x_1 = x_0 - \frac{g(x_0)}{g'(x_0)}$$

Then $|r - x_1| \leq M |r - x_0|^2$ where $M = \sup_{\xi \in (r, x_0)} \left| \frac{g''(\xi)}{2 g'(x_0)} \right|$

Consequence: $|r - x_2| \leq M |r - x_1|^2 \leq M^3 |r - x_0|^4$

If $|r - x_0| \leq \frac{1}{2}$ $\quad M < 1$ then to get to $|r - x_t| \leq \varepsilon$ we need $O\left(\lg \lg \frac{1}{\varepsilon}\right)$ many iterations.

$$|r - x_t| \leq |r - x_0|^{2^t} \leq 2^{-2^t} \leq \varepsilon \iff \frac{1}{\varepsilon} \leq 2^{2^t}$$

$$\iff \lg \lg \frac{1}{\varepsilon} \leq t$$

$$|r - x_0| \le 2^{-2^t} \le \varepsilon \iff \frac{1}{\varepsilon} \le 2^{2^t}$$

$$\iff \log \log \frac{1}{\varepsilon} \le t$$

Pf: Second order expansion & MVT.

MVT: $h : \mathbb{R} \to \mathbb{R}$ continuous on $[a, b]$ & diff on (a, b) then $h'(c) = \dfrac{h(b) - h(a)}{b - a}$

① $\Rightarrow$ $g(r) = g(x_0) + g'(x_0)(r - x_0) + \frac{1}{2} g''(\xi)(r - x_0)^2$

for $\xi \in (r, x_0)$ ($\because g''$ is continuous)

we know $x_1 = x_0 - \dfrac{g(x_0)}{g'(x_0)}$

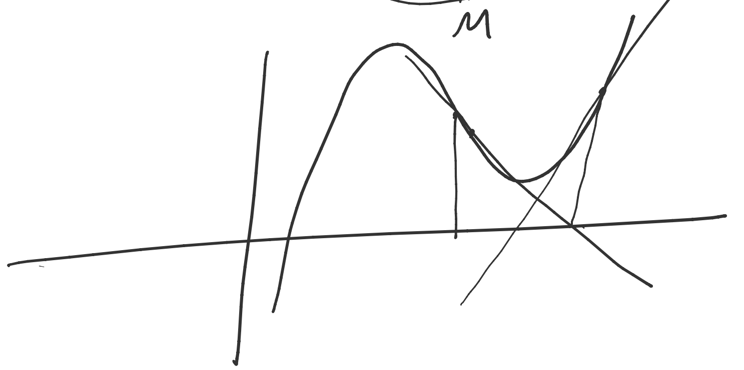
$$\Rightarrow g(x_0) = g'(x_0)(x_0 - x_1)$$

① $\Rightarrow$ $0 = g'(x_0)(x_0 - x_1) + g'(x_0)(r - x_0) + \frac{1}{2} g''(\xi)(r - x_0)^2$

$\iff$ $g'(x_0)(r - x_1) + \frac{1}{2} g''(\xi)(r - x_0)^2 = 0$

$\iff$ $|g'(x_0)(r - x_1)| = \left|\frac{1}{2} g''(\xi)(r - x_0)^2\right|$

$$|r - x_1| \le \Bigg|\underbrace{\frac{\frac{1}{2} g''(\xi)}{g'(x_0)}}_{M}\Bigg| \, |r - x_0|^2$$



- Multivariate (root finding)

$g : \mathbb{R}^n \to \mathbb{R}^n$

$g_1(x_1, \dots, x_n)$

$\vdots$

$g_n(x_1, \dots, x_n)$

$$x_{t+1} = x_t - \frac{g(x_t)}{g'(x_t)}$$

$$J_g = \left[\frac{\partial g_i}{\partial x_j}\right]_{1 \le i, j \le n}$$

$g_n(x_1, \ldots, x_n)$ $\quad \Big]_{1 \le i, j \le n}$

$$x_{t+1} = x_t - \left(J_g(x_t)\right)^{-1} \cdot g(x_t)$$

- Unconstrained Optimization

f is convex, twice differentiable.

$$x^* = \arg\min_x f(x)$$

$$\Updownarrow$$

$$\nabla f(x) = 0 \qquad \nabla f(x) \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]$$

$$g = \nabla f : \mathbb{R}^n \to \mathbb{R}^n.$$

$$x_{t+1} = x_t - J_g(x_t)^{-1} g(x_t)$$

$$= x_t - \underbrace{J_{\nabla f}(x_t)^{-1}}_{\left[\nabla^2 f(x_t)\right]^{-1}} \nabla f(x_t)$$

$$x_{t+1} = x_t - \ {}^2 f(x_t)\big]^{-1} \cdot \nabla f(x_t)$$

Alternate View:

$$\min_y f(y) = \underline{f(x)} + \nabla f(x)(\ \ - x) + \frac{1}{2}\underbrace{\nabla^2 f(x)(y-x)^2}_{(y-x)^T \nabla^2 f(x)(y-x)}$$

$$\nabla_y f = 0 \iff \nabla f(x) + \frac{\cancel{2}}{\cancel{2}} \nabla^2 f(x)(y - x) = 0$$

$$y - x = \left[\nabla^2 f(x)\right]^{-1}(-\nabla f(x))$$

$$\left[\nabla^2 f(x)\right]^{-1} \nabla f(x)$$

$$\Rightarrow y = x - \left[\nabla^2 f(x)\right]$$

☆ Newton's Decrement:

$$\lambda(x) = \sqrt{\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)} \leftarrow \|\nabla f(x)\|_{\nabla^2 f(x)^{-1}}$$

☆ Newton's De [illegible] ← || [illegible]

$$\lambda_f(x) = \sqrt{\nabla f(x)^T \, \nabla^2 f(x)^{-1} \nabla f(x)}$$

$$\boxed{\text{at } x^*, \ \nabla f(x^*) = 0 \ \Rightarrow \ \lambda_f(x^*) = 0}$$

$$\propto \ f(x) - \min_y \left(f(x) + \nabla f(x)(y-x) + \frac{1}{2} \nabla^2 f(x)(y-x)^2 \right)$$

$$= \cancel{f(x)} - \cancel{f(x)} - \nabla f(x)\left(\cancel{x} - \left[\nabla^2 f(x)\right]^{-1} \nabla f(x) - \cancel{x}\right) - \frac{1}{2}\left[\nabla^2 f(x)\right]\left(\cancel{x} - \left[\nabla^2 f(x)\right]^{-1} \nabla f(x) - \cancel{x}\right)^2$$

$$= \underline{\nabla f(x)^T \left[\nabla^2 f(x)\right]^{-1} \nabla f(x)} - \frac{1}{2}\left[\underbrace{\nabla f(x)^T \underline{\left[\nabla^2 f(x)\right]^{-1}} \cancel{\left[\nabla^2 f(x)\right]} \cancel{\left[\nabla^2 f(x)\right]^{-1}} \nabla f(x)}\right]$$

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

$$= \frac{\lambda_f(x)^2}{2}$$

Thm (quadratic decrease):

If for all x, $\left|\nabla^3 f(x)\,[h,h,h]\right| \leq 2\left(h^T \nabla^2 f(x) h\right)^{3/2} \quad \forall h$

$$\nabla^3 f(x) = \left[T_{ijk}\right]_{1 \leq i,j,k \leq n} [h,h,h] = \sum_{i,j,k} T_{ijk} h_i h_j h_k$$

Then, $\lambda_f\big(\underbrace{x - \left[\nabla^2 f(x)\right]^{-1} \nabla f(x)}_{x'}\big) \leq \left(\dfrac{\lambda_f(x)}{1 - \lambda_f(x)}\right)^2$