Week 6

OPTIMALITY CONDITIONS

Consider a simple optimization problem of form,

$$\min f(x)$$

$$x \in \mathcal{X} = \{x \mid f_i(x) \leq 0, Ax = b\}$$

Assume that this optimization problem is convex.

What properties does x^* satisfy?

Unconstrained : $\nabla f(x^*) = 0$

First order convexity condition : $f(x) \geq f(y) + \langle \nabla f(y), x-y \rangle$,

$x, y \in \mathcal{X}$ (feasible).

Suppose that $\langle \nabla f(y), x-y \rangle \geq 0 \quad \forall x \in \mathcal{X}$, for some y.

then $f(x) \geq f(y) \quad \forall x \in \mathcal{X}$

(or)

$$y = \arg\min_{x \in \mathcal{X}} f(x) = x^*.$$

General optimality condition for general optimization problems (convex optimization problems)

$$\boxed{\langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad , \forall x \in \mathcal{X}}$$

Example.

If $\mathcal{X} = \mathbb{R}^n$, then $\langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in \mathbb{R}^n$.

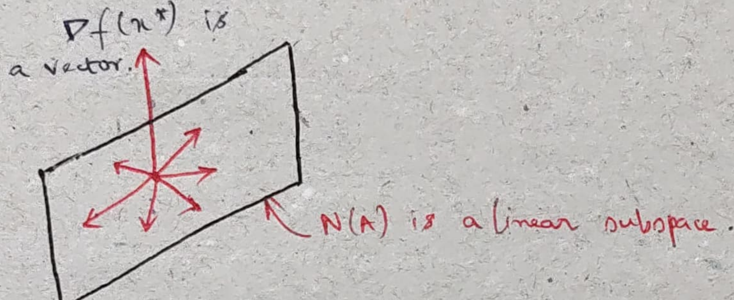
$$\Rightarrow \nabla f(x^*) = 0.$$

If $\min_{x \in \mathcal{X} = \{x \mid Ax = b\}} f(x)$, then $\langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in \mathcal{X}$.

Note:

$$\left.\begin{array}{l} x \in \mathcal{X} \Rightarrow Ax = b \\ x^* \in \mathcal{X} \Rightarrow Ax^* = b \end{array}\right\} \Rightarrow A(x - x^*) = 0$$

(or) $x - x^* \in N(A) \quad \forall x \in \mathcal{X}$.



$$\langle \nabla f(x^*), u \rangle \geq 0, \quad u \in N(A).$$

$$\Rightarrow \langle \nabla f(x^*), x - x^* \rangle = 0.$$

(ie) $\nabla f(x^*)$ and $x - x^*$ makes $90°$ angle.

$$\Rightarrow \nabla f(x^*) \in N(A)^{\perp} = R(A^T)$$

$$\text{or} \quad \exists v : \nabla f(x^*) = A^T v$$

Thus, the optimality condition for Affine equality constraints is $\nabla f(x^*) \in R(A^T)$.

Example:

$$\min_{x_1, x_2} f_1(x_1) + f_2(x_2)$$

$$\text{s.t. } x_1 + x_2 = 1$$

$x \in \mathbb{R}^2$

$A = [1 \ \ 1]$, $A^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$$\nabla f(x) = \begin{bmatrix} f_1'(x_1) \\ f_2'(x_2) \end{bmatrix}$$

$$\begin{bmatrix} f_1'(x_1^*) \\ f_2'(x_2^*) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} v \quad , \; v \in \mathbb{R}.$$

So, the optimality conditions are,

(1) $f_1'(x_1^*) = f_2'(x_2^*) = v$ (first order condition)

(2) $x_1^* + x_2^* = 1$ (feasibility)

So, this is the optimality conditions in the general case.

Next, lets look into Lagrange duality which provides a very principled way of deriving all sorts of optimality conditions, especially for convex optimization problems, though it is applicable for both convex and non-convex optimization problems.

Lagrange Duality

- Allows us to go beyond just "Solving" problems
- Allows bounding techniques
- Allows us to obtain closed-form solution
- We are able to develop good algorithms using duality theories.

Consider

$$x^* = \arg\min f_0(x)$$
$$\text{s.t. } f_i(x) \le 0, \; i = 1 \dots m$$
$$h_j(x) = 0, \; j = 1 \dots p$$
$$x \in \mathcal{D} \text{ (domain)}$$

$$\left.\begin{matrix} \lambda_1, \lambda_2, \dots, \lambda_m \\ \nu_1, \nu_2, \dots, \nu_p \end{matrix}\right\} \text{Dual variables}$$

$$p = f_0(x^*) = \begin{cases} \infty & \text{(infeasible)} \\ -\infty & \text{(unbounded below)} \\ \text{finite} & \text{(otherwise)} \end{cases}$$

We will associate some variables with different constraints.

Lagrange Multipliers

$$L(\underline{x}, \underline{\lambda}, \underline{\nu}) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) + \sum_{j=1}^{p} \nu_j h_j(x).$$

where, $\underline{x} \to \mathbb{R}^n$, $\underline{\lambda} \to \mathbb{R}^m$, $\underline{\nu} \to \mathbb{R}^p$

$x \to$ Primal variable

$\lambda, \nu \to$ Dual variables.

- It is Affine in λ, ν
- When the problem is convex, $\lambda_i \geq 0$.
  - f_0 convex
  - f_i convex
  - h_j affine

  $\rightarrow L(x, \lambda, \nu)$ is convex in x.

Dual function

$$g(\lambda, \nu) = \min_{x \in \mathcal{D}} L(x, \lambda, \nu)$$

$g(\lambda, \nu)$ is concave in λ, ν

This is a pointwise minimum of affine functions.

(by definition, even when original problem is not convex).

Consider : $\tilde{x}$ which is feasible $\Rightarrow \tilde{x} \in \mathcal{D}$ (Domain of Primal function)

$$f_i(\tilde{x}) \leq 0, \ i = 1 \dots m$$
$$h_j(\tilde{x}) = 0, \ j = 1 \dots p$$

$$f_0(\tilde{x}) \geq f_0(\tilde{x}) + \overbrace{\sum_{i=1}^{m} \underbrace{\lambda_i}_{+ve} \underbrace{f_i(\tilde{x})}_{-ve}}^{-ve} + \cancel{\sum_{j=1}^{p} \nu_j \underbrace{h_j(\tilde{x})}_{0}} \rightarrow 0.$$

$$= L(\tilde{x}, \lambda, \nu), \quad \nu \in \mathbb{R}^p, \ \lambda \in \mathbb{R}^m_+$$

$$\geq \min_{x \in \mathcal{D}} L(x, \lambda, \nu) \rightarrow \text{By definition of minimum operator}$$

$$= g(\lambda, \nu) \rightarrow \text{Dual function}$$

$$f_0(\tilde{x}) \geq g(\lambda, \nu) \quad \text{for all } \lambda \geq 0, \ \tilde{x} \text{ feasible}$$

Primal objective — Dual objective

Weak Duality

Recall, $f_0(\tilde{x}) \geq g(\lambda, \nu)$, $\lambda \geq 0$, $\tilde{x}$ feasible.

Dual problem is defined as,

$$D = \max_{\lambda \geq 0} g(\lambda, \nu) = -\min -g(\lambda, \nu)$$

concave

$(\lambda, \nu) \in \text{dom}(g)$ (implicit constraint),

Primal Problem is defined as,

$$P = \min f_0(x)$$

$f_i(x) \leq 0, \; i = 1 \dots m$

$h_j(x) = 0, \; j = 1 \dots p$

$x \in \mathcal{D}$

Now, the inequality

$$f_0(\tilde{x}) \geq g(\lambda, \nu)$$

depends on $\tilde{x}$: $f_i(\tilde{x}) \leq 0$, $h_j(\tilde{x}) = 0$, $\tilde{x} \in \mathcal{D}$

depends on λ, ν: $\lambda \geq 0$, $(\lambda, \nu) \in \text{dom}(g)$

$\tilde{x} \to$ Primal feasible

$\lambda \to$ Dual feasible

So, we can tighten this inequality as below.

$$\Rightarrow P = \min_{\tilde{x}} f_0(\tilde{x}) \geq \max g(\lambda, \nu) = D.$$

$f_i(\tilde{x}) \leq 0$ $\quad$ $\lambda \geq 0$

$h_j(\tilde{x}) = 0$ $\quad$ $(\lambda, \nu) \in \text{dom}(g)$

$\tilde{x} \in \mathcal{D}$

$\Rightarrow \boxed{P \geq D} \longrightarrow$ Weak Duality, by definition.

Note: Suppose primal is unbounded below.

(1) $P = -\infty \geq D \Rightarrow D = -\infty$

$= g(\lambda^*, \nu^*) = -\infty$ (or)

$= -\min_{\lambda \geq 0} -g(\lambda, \nu)$

$\Rightarrow$ Dual problem is infeasible.

(2) $D = \infty$ (dual is unbounded above)

$\Rightarrow P = \infty \Rightarrow$ Primal infeasible.

$$\boxed{\text{Duality Gap} : P - D \geq 0}$$

Example. $\min \frac{1}{2} x^T x$ | $b \in \mathbb{R}^p$, $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{p \times n}$ | $\nu = \begin{bmatrix} \nu_1 \\ \vdots \\ \nu_p \end{bmatrix}$

$Ax = b$ ...... $\nu \in \mathbb{R}^p$

This is the least norm solution.

consider $\text{rank}(A) = p < n$.

$$L(x, \nu) = \frac{1}{2} x^T x + \nu^T (Ax - b)$$

$$g(\nu) = \min_x L(x, \nu) = \min_x \frac{1}{2} x^T x + \nu^T (Ax - b)$$

Solution is given by,

$$\nabla_x L(x, \nu) = 0$$

(or) $x + A^T \nu = 0$

(or) $x = -A^T \nu$.

so, $g(\nu) = \frac{1}{2} \nu^T A^T A \nu + \nu^T(-AA^T \nu - b)$

$= -\frac{1}{2} \nu^T AA^T \nu - b^T \nu$

Dual Problem : $D = \max -\frac{1}{2} \nu^T AA^T \nu - b^T \nu$, $\nu \in \mathbb{R}^p$

$\nabla_\nu g(\nu) = 0 \Rightarrow -AA^T \nu - b = 0$

(or) $\nu^* = -(AA^T)^{-1} b$

$D = g(\nu^*) = \frac{1}{2} b^T (AA^T)^{-1} b$.

Therefore, $\frac{1}{2} \tilde{x}^T \tilde{x} \geq \frac{1}{2} b^T (AA^T)^{-1} b$

for any $\tilde{x}$: $A\tilde{x} = b$

Also, $\boxed{P \geq \frac{1}{2} b^T (AA^T)^{-1} b}$

Examples

Let us look into various examples/usefulness of weak duality.

⊙ Dual formulation and the Dual problem is formulation dependent. Different ways of taking the Dual will result in different dual problems. Let us consider the problem,

$$\min_x - \sum_{i=1}^m \log(b_i - a_i^T x).$$

This is an unconstrained problem.

$$= \min_{(x, y)} \; - \sum_{i=1}^{m} \log(y_i)$$

$$\text{s.t. } y = b - Ax$$

$y_i = b_i - a_i^T x$

(or) $y = \underline{b} - Ax$ (in compact)

Lagrangian,

$$\underbrace{L(x, y}_{\text{Primal variables}}, \underbrace{\nu}_{\text{Dual variable}}) = -\sum_{i=1}^{m} \log(y_i) + \nu^T(y - b + Ax)$$

$$= \sum_{i=1}^{m} \log(y_i) + \sum \nu_i (y_i - b_i + a_i^T x)$$

$$= \underbrace{\sum_{i=1}^{m} (-\log y_i + \nu_i y_i)}_{\text{depends on } y_i} + \underbrace{\sum_{i=1}^{m} \nu_i (a_i^T x)}_{\text{depends on } x} - \underbrace{\sum_{i=1}^{m} \nu_i b_i}_{\text{constant w.r.t } x, y}$$

Dual function,

$$g(\nu) = \min_{x, y} L(x, y, \nu)$$

$$= \min_{y > 0} \sum_{i=1}^{m} (\nu_i y_i - \log y_i) + \min_{x} \sum_{i=1}^{m} \nu_i (a_i^T x) - \nu^T b$$

$$= \underbrace{\sum_{i=1}^{m} \min_{y_i > 0} (\nu_i y_i - \log y_i)}_{①} + \underbrace{\min_{x} \left(\sum_{i=1}^{m} a_i \nu_i\right)^T x}_{②} - \underbrace{\nu^T b}_{③}$$

① $\Rightarrow$ $\dfrac{d}{dy_i}(\;) = 0$

$$\nu_i - \frac{1}{y_i} = 0$$

(or) $y_i = \dfrac{1}{\nu_i}$ (only valid when $\nu_i > 0$)

Note : when $\nu_i \leq 0$ then

$\nu_i y_i - \log(y_i) \rightarrow -\infty$ (unbounded below)

So,

- if $\nu_i > 0$, then objective $= 1 + \log(\nu_i)$
- if $\nu_i < 0$, then objective $\rightarrow -\infty$

So, $$\min_{y_i \geq 0} \nu_i y_i - \log(y_i) = \begin{cases} 1 + \log(\nu_i) & , \nu_i > 0 \\ -\infty & , \nu_i \leq 0 \end{cases} \quad \Bigg| \; i = 1, \ldots, m$$

② $\Rightarrow \min_x \underbrace{\left(\sum a_i \nu_i\right)^T}_{c} x = \min_x c^T x = \begin{cases} 0, & c = 0 \\ -\infty, & \text{otherwise} \end{cases}$

$$c = \begin{cases} 0, & \text{when } \sum a_i \nu_i = 0 \quad (\text{or}) \quad A^T \nu = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

Summary.

$$g(\nu) = \begin{cases} \sum_{i=1}^{m} (1 + \log(\nu_i)) + 0 - b^T \nu, & \text{if } \nu \geq 0,\ A^T \nu = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

(or)

$$-g(\nu) = \begin{cases} -\sum_{i=1}^{m} (1 + \log(\nu_i)) + b^T \nu, & A^T \nu = 0,\ \nu \geq 0 \\ \infty, & A^T \nu \neq 0,\ \nu \leq 0 \end{cases}$$

Recall : Extended function value definition
(convex function $= \infty$ outside domain)

$$\Rightarrow \operatorname{dom}(g) = \{\nu \mid -g(\nu) < \infty\}$$
$$= \{\nu \mid A^T \nu = 0,\ \nu \geq 0\}$$

Dual Problem :

$$\max_\nu \quad m + \sum_{i=1}^{m} \log(\nu_i) - b^T \nu$$
$$\text{s.t.} \quad \left.\begin{array}{l} \nu \geq 0 \\ A^T \nu = 0 \end{array}\right\} \leftarrow \text{Domain restrictions. } \nu \in \operatorname{dom}(g) \text{ (implicit)}$$

Summary :

- Dual is formulation-dependent
- Pay attention to domains.

Integer Programming dual

We may not dualize all the constraints.

Integer Program : $\min c^T x$ $\quad \Big| \; x \in \{0,1\}^n$
$\text{s.t. } Ax \leq b$ $\quad \Big| \; x \in \mathbb{R}^n$
$x_i \in \{0,1\}$ ← Not dualize

Solving an integer program is generally known to be NP-hard.
(a) time $= O(e^n)$.

---

define $f(x) = \{c^T x \mid x_i \in \{0,1\}\}$

$$\mathrm{dom}(f) = \{0,1\}^n .$$

$$\Rightarrow \min_x f(x)$$

$$Ax \leq b \;\ldots\; \lambda \in \mathbb{R}^m$$

This objective is non-convex.

Lagrangian,

$$L(x, \lambda) = c^T x + \lambda^T (Ax - b)$$

$$\mathrm{dom}(L) = \underbrace{\{0,1\}^n}_{x} \times \underbrace{\mathbb{R}^m}_{\lambda}$$

$$= \mathbb{B}^n \times \mathbb{R}^m$$

$\mathbb{B}^n \rightarrow$ Binary vector

$\mathbb{B}^n = \{0,1\}^n$.

The Dual function,

$$g(\lambda) = \min_{x \in \mathbb{B}^n} L(x, \lambda) = \min_{x \in \mathbb{B}^n} c^T x + \lambda^T (Ax - b)$$

$$= \min_{x \in \mathbb{B}^n} (c + A^T \lambda)^T x - \lambda^T b$$

$$= \min_{x_i \in \{0,1\}} \sum_{i=1}^{n} [c + A^T \lambda]_i \, x_i - \lambda^T b$$

$$= \sum_{i=1}^{n} \min_{x_i \in \{0,1\}} [c + A^T \lambda]_i \, x_i - \lambda^T b$$

Aside:

$$\min_{x_i \in \{0,1\}} h_i x_i$$

pick $x_i^* = \begin{cases} 0, & h_i \geq 0 \\ 1, & h_i \leq 0 \end{cases}$

so, $h_i x_i^* = \begin{cases} 0, & h_i \geq 0 \\ h_i, & h_i \leq 0 \end{cases} = \min[h_i, 0]$

so, $g(\lambda) = \sum_{i=1}^{m} \min\left([c + A^T\lambda]_i, 0\right) - \lambda^T b$

Dual Problem : $\max_{\lambda \geq 0} g(\lambda)$

$$= \max_{\lambda \geq 0} \sum_{i=1}^{m} \min\left\{[c + A^T\lambda]_i, 0\right\} - \lambda^T b.$$

Using Epigraph trick,

$$= \max_{t, \lambda} \sum_{i=1}^{m} t_i - \lambda^T b$$

when $\lambda \geq 0$,

$$= \min\left([c + A^T\lambda]_i, 0\right) \geq t_i \quad , i = 1, \dots, m.$$

Dual of IP :

$$D = \max_{t, \lambda} \mathbf{1}^T t - b^T \lambda$$

$$c + A^T\lambda \geq t$$

$$t \leq 0$$

This is a LP.

From weak duality : $P \geq D$ (D can be easily found)

This is practically used a lot for solving IP.

Take away :

- Some constraints may be included in dom.
- Need not dualize them
- get different dual problems.

Karush-Kuhn-Tucker (KKT) conditions

So far we've seen Lagrangian duality and weak duality which are used for general applications, and various implications of that result. We also derived the dual of some example problems such as Integer programming problem. Now, lets look into KKT conditions which are specifically for convex optimization.

Assume : ① x^* is the primal optimum
② (λ^*, ν^*) is the dual optimum
③ $P = D$, which is strong duality

Then the KKT conditions hold.

Dual optimum is,

$$D = g(\lambda^*, \nu^*) = \min_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*) \quad \text{(definition)}$$

Should be equality

$$\overset{①}{\leq} L(x^*, \lambda^*, \nu^*) \quad \text{(should be equality)}$$

$$= f_0(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i^* f_i(x^*)}_{\text{② should be 0}} + \underbrace{\sum_{j=1}^{p} \nu_j^* h_j(x^*)}_{\text{③ should be 0}}$$

$$\boxed{= f_0(x^*) = P}$$ ← Should hold from strong duality.

$P = D$ from strong duality.

KKT conditions

1. Primal feasibility $f_i(x^*) \leq 0$, $h_j(x^*) = 0$
2. Dual feasibility $\lambda_i^* \geq 0$
3. Complementary slackness : $\lambda_i^* f_i(x^*) = 0$

   $\Rightarrow$ either $\lambda_i = 0$ (or) $\underbrace{f_i(x^*) = 0}$ (or) both

   ↑ x^* on the boundary of f_i

   $\|x_i\|^* - 1 \leq 0$

4. Stationary condition

$$x^* = \arg\min_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*)$$

Under these 4 conditions; If $P = D$, x^* and λ^* exists, then it follows KKT conditions.

Unconstrained case : $\nabla_x L(x^*, \lambda^*, \nu^*)\Big|_{x = x^*} = 0.$

$(\mathcal{D} = \mathbb{R}^n)$

for convex problems

Summarize:

① Optimum (x^*, λ^*, ν^*), $P = D$ $\iff$ KKT conditions

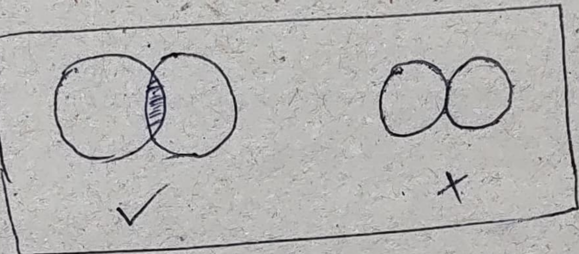
Note: convex $\not\Rightarrow$ KKT or $P = D$ ...

Slater's Theorem

f_0, f_i convex
h_{ij} affine
P finite
$\underbrace{\text{Slater's Constraint Qualification (CQ)}}$

$\Rightarrow$ $P = D$, λ^* finite, KKT conditions.

There exists $\tilde{x}$ s.t. $f_i(\tilde{x}) < 0 \ \forall i = 1, \dots, m$ for non-linear f_i



Example: $P_1 = \min_x \frac{1}{2} x^T P x + q^T x$ ← Convex $(P \succeq 0)$

s.t. $Ax = b$ $(f_i = 0)$

$b \in \mathbb{R}(A) \Rightarrow$ feasible + not unbounded below $\Rightarrow$ P_1 is finite.

$\Rightarrow P = D$ & ① $Ax^* = b$

② $\nabla L(x, \nu) = 0 = Px^* + q + A^T \nu^*$

$$L(x, \nu) = \frac{1}{2} x^T P x + q^T x + \nu^T (Ax - b)$$

$$\Rightarrow \nabla \left(\frac{1}{2} x^T P x + q^T x + \nu^T (Ax - b) \right) = 0$$

Solve KKT? $\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$

$\Rightarrow$ Solve for (x^*, ν^*).

Water filling example

Demonstration of how KKT conditions help us understand the problem better.

Equalization of communication channels

$$\max \text{ (sum rate)}$$
$$\text{s.t. total power} \leq \text{budget}$$

(Assume No ICI)

i.e. we are allocating power to each of the channels, maximizing the sum rate, so that the total power that is allocated to all this bunch of channels is within the limits. This kind of setting is common in OFDMA, TDMA, ZF-MIMO, ...

Rate of i^{th} channel,

$$\text{rate}_i = \log(1 + SNR_i)$$
$$= \log(1 + \gamma_i p_i)$$

Allocated power

$$\gamma_i = \frac{g_i}{N_0} \geq 0$$

Channel gain

Noise

The optimization problem becomes

$$\max_{\underline{p}} \sum_{i=1}^{n} \log(1 + \gamma_i p_i)$$
$$\text{s.t. } \sum_{i=1}^{n} p_i \leq P$$
$$p_i \geq 0$$

Power budget

Note:

- It is clearly convex, as log is concave function
- Affine constraints (Slater's Constraint Qualification not needed)
- Feasible (eg. $p_i = 0 \ \forall \ i$)
- Not unbounded below.

$P = D$ & KKT

- There are different ways of writing the Dual
- Let us keep $p_i \geq 0$ implicit (not dualize)
  - Dual problem would be one dimensional (scalar) problem

$$\min_{p \geq 0} - \sum_{i=1}^{n} \log(1 + p_i \gamma_i) = \min_{\sum p_i \leq P} f(\underline{p})$$
$$\text{s.t. } \sum_{i=1}^{n} p_i \leq P, \ p_i \geq 0$$

objective : $f(p) = -\sum_{i=1}^{n} \log(1+p_i\gamma_i)$

but, $\text{dom}(f) = \mathbb{R}_+^n$

KKT condition.

(a) Stationarity.

$$p^* = \arg\min_{p \geq 0} L(p, \lambda) = -\sum \log(1+p_i\gamma_i) + \lambda^*\left(\sum p_i - P\right)$$

$$p^* = \arg\min_{p \geq 0} \sum_{i=1}^{n} \left[-\log(1+p_i\gamma_i) + \lambda^* p_i\right]$$

Split into n sub problems.

$$\Rightarrow p_i^* = \arg\min_{p_i \geq 0} -\log(1+p_i\gamma_i) + \lambda^* p_i$$

Case (a): Suppose $\lambda^* = 0$, then

$$p_i^* = \arg\min_{p_i \geq 0} \underbrace{-\log(1+p_i\gamma_i)}_{\longrightarrow -\infty \text{ as } p_i^* \to \infty}$$

so, unbounded below and $\sum p_i^* \leq P$

$\Rightarrow \lambda^* > 0$ (or) $\lambda^* \neq 0$.

Case (b): $\lambda^* > 0$

- b(i) Suppose $p_i^* > 0$,

$$\frac{d}{dp_i}\left(-\log(1+p_i\gamma_i) + \lambda^* p_i\right) = 0$$

(or) $\lambda^* = \dfrac{\gamma_i}{1+p_i^*\gamma_i}$ (or) $\boxed{p_i^* = \dfrac{1}{\lambda^*} - \dfrac{1}{\gamma_i} > 0}$

This is the solution.

- b(ii) Suppose $p_i^* = 0$

In this case, $\dfrac{1}{\lambda^*} - \dfrac{1}{\gamma_i} \leq 0$.

Therefore,

(a) $p_i^* = \max\left\{0, \frac{1}{\lambda^*} - \frac{1}{\gamma_i}\right\} = \left[\frac{1}{\lambda^*} - \frac{1}{\gamma_i}\right]_+$

KKT

(b) $p^* \geq 0, \; \sum p_i^* \leq P$

(c) $\lambda^* > 0, \; \lambda^* > 0$

(d) $\lambda^*\left(\sum_{i=1}^{n} p_i^* - P\right) = 0 \Rightarrow \sum p_i^* = P$

solve KKT ?

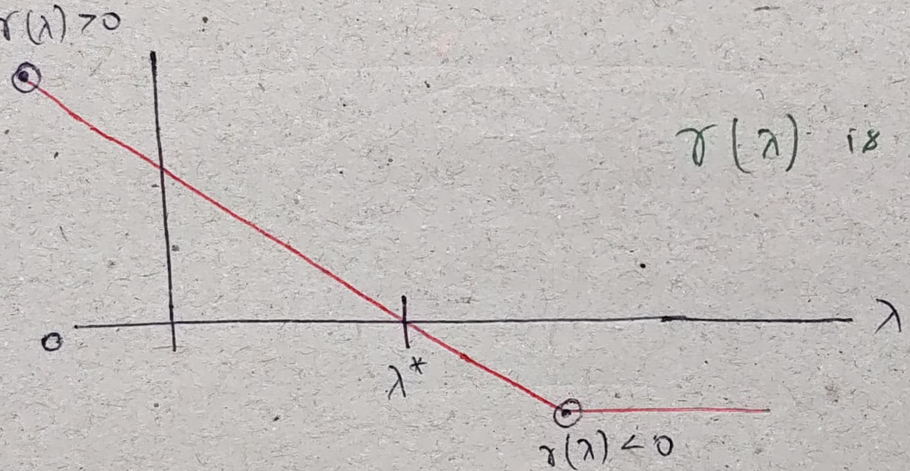
$$\gamma(\lambda) = \sum_{i=1}^{n} \left[\frac{1}{\lambda} - \frac{1}{\gamma_i^*}\right]_+ - P = 0$$

λ^* is the root of $\gamma(\lambda)$.

Intuition :

λ is small $\Rightarrow \gamma(\lambda)$ is large

λ is large $\Rightarrow \gamma(\lambda)$ is small



$\gamma(\lambda)$ is decreasing.

Bisection Algorithm

$\gamma(\lambda)$ is decreasing

while $\lambda_{max} - \lambda_{min} > \varepsilon$ (epsilon)

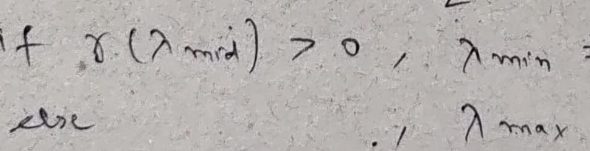
$$\lambda_{mid} = \frac{\lambda_{min} + \lambda_{mid}}{2}$$

if $\gamma(\lambda_{mid}) > 0, \; \lambda_{min} = \lambda_{mid}$

else $, \; \lambda_{max} = \lambda_{mid}$

$\lambda \in [\lambda_{min}, \lambda_{max}]$

$B = \lambda_{max} - \lambda_{min}$



k steps

$$\lambda_{max} - \lambda_{min} \sim \frac{B}{2^k} = \varepsilon$$

No. of iterations

when $k = \log_2\left(\frac{B}{\varepsilon}\right)$

And, this is called iteration complexity.

Eg $B = 1$, $\varepsilon = 10^{-3}$, $k \approx 10$.

$\varepsilon = 10^{-6}$, $k \sim 20$

⋮