Exercice n° 1 :

Assume that $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^{m}$

$$(\mathcal{P}) : \begin{cases} \min_{x} & \|x\|_\infty \\ \text{s.t.} & Ax = b \end{cases} \iff \begin{cases} \min_{x,t} & t \\ \text{s.t.} & \|x\|_\infty \leqslant t \\ & Ax = b \\ & t \geqslant 0 \end{cases}$$

• We notice that : $\|x\|_\infty \leqslant t \iff \forall i \in [\![1,n]\!],\ |x_i| \leqslant t$
(for $t \geqslant 0$)

$\iff \forall i \in [\![1,n]\!],\ x_i \leqslant t$ and $-x_i \leqslant t$

$\iff x \leqslant t\,\mathbb{1}_n$ and $-x \leqslant t\,\mathbb{1}_n$

(where $\mathbb{1}_n = [1, \dots, 1]^T \in \mathbb{R}^n$)

Thus :

$$\boxed{(\mathcal{P}) \iff \begin{cases} \min_{x,t} & t \\ \text{s.t.} & Ax = b \\ & t \geqslant 0 \\ & x \leqslant t\,\mathbb{1}_n \\ & -x \leqslant t\,\mathbb{1}_n \end{cases}}$$

• The Lagrangian associated to $(\mathcal{P})$:

Let $x \in \mathbb{R}^n$, $\nu \in \mathbb{R}^m$, $\alpha \in \mathbb{R}$, $t \in \mathbb{R}$, $\beta \in \mathbb{R}^n$, $\gamma \in \mathbb{R}^n$.

$$\mathcal{L}(x, t, \alpha, \beta, \gamma, \nu) = t - \alpha t + \beta^T(x - t\,\mathbb{1}_n) + \gamma^T(-x - t\,\mathbb{1}_n) + \nu^T(Ax - b)$$

$$= \left[1 - \alpha - \mathbb{1}_n^T(\beta + \gamma)\right] t + \left[A^T\nu + \beta - \gamma\right]^T x - b^T\nu$$

$$\Rightarrow g(\alpha, \beta, \gamma, \nu) = \inf_{x,t} \mathcal{L}(x, t, \alpha, \beta, \gamma, \nu)$$

$$= \inf_{t} \left(1 - \alpha - \mathbb{1}_n^T(\beta + \gamma)\right) t + \inf_{x} \left[A^T\nu + \beta - \gamma\right]^T x - b^T\nu$$

(because $\mathcal{L}$ is a sum of separable functions of x and t).

• $\inf_{t \in \mathbb{R}} \left(1 - \alpha - \mathbb{1}_n^T(\beta + \gamma)\right) t = \begin{cases} 0 & \text{if } 1 - \alpha - \mathbb{1}_n^T(\beta + \gamma) = 0 \\ -\infty & \text{otherwise} \end{cases}$

• $\inf_{x \in \mathbb{R}^n} \left[A^T\nu + \beta - \gamma\right]^T x = \begin{cases} 0 & \text{if } A^T\nu + \beta - \gamma = 0 \\ -\infty & \text{otherwise} \end{cases}$

2

Therefore, $g(\alpha,\beta,\gamma,\nu) = \begin{cases} -b^T\nu & \text{if } 1-\alpha-\mathbb{1}_m^T(\beta+\gamma)=0 \text{ and } A^T\nu+\beta-\gamma=0 \\ -\infty & \text{otherwise.} \end{cases}$

Then, the dual is given by:

$$(D): \begin{cases} \max & g(\alpha,\beta,\gamma,\nu) \\ \text{st} & \alpha \geq 0 \\ & \beta \geq 0 \\ & \gamma \geq 0 \end{cases} \iff \begin{cases} \max & -b^T\nu \\ \text{s.t.} & \alpha \geq 0 \\ & \beta \geq 0, \quad \gamma \geq 0 \\ & \mathbb{1}_m^T(\beta+\gamma) = 1-\alpha \\ & A^T\nu+\beta-\gamma=0 \end{cases}$$

$$\iff \boxed{\begin{cases} \max & -b^T\nu \\ \text{s.t.} & \alpha \geq 0 \\ & \beta \geq 0, \ \gamma \geq 0 \\ & \mathbb{1}_m^T(\beta+\gamma) \leq 1 \\ & A^T\nu+\beta-\gamma=0 \end{cases}}$$

Exercice n°2:

For $u, v \in \mathbb{R}^n_{++}$, $\quad D_{KL}(u,v) := \sum_{i=1}^{n} u_i \log\left(\frac{u_i}{v_i}\right) - u_i + v_i$

Let $f : x \in \mathbb{R}^n_{++} \longmapsto \sum_{i=1}^{n} x_i \log(x_i)$ and $\varphi : t \in \mathbb{R}_{++} \longmapsto t \log(t)$

- φ is twice-differentiable and $\forall t > 0$, $\varphi''(t) = \frac{1}{t} > 0$

  Then, φ is strictly convex over $\mathbb{R}_{++}$.

- We notice that for $x \in \mathbb{R}^n_{++}$, $f(x) = \sum_{i=1}^{n} \varphi(x_i)$

  f is a sum of strictly convex functions. So, f is strictly convex on $\mathbb{R}^n_{++}$.

  In fact: $\forall\ x, y \in \mathbb{R}^n_{++}$, $\forall \alpha \in [0,1]$:

  $$f(\alpha x + (1-\alpha)y) = \sum_{i=1}^{n} \varphi(\alpha x_i + (1-\alpha) y_i) < \sum_{i=1}^{n} \alpha \varphi(x_i) + (1-\alpha)\varphi(y_i) = \alpha f(x) + (1-\alpha) f(y)$$

- The gradient of f is given by: (for $v \in \mathbb{R}^n_{++}$)

  $$\nabla f(v) = \left[1 + \log(v_1), \dots, 1 + \log(v_n)\right]^T \in \mathbb{R}^n$$

- $\forall\ u, v \in \mathbb{R}^n_{++}$:

$$\begin{aligned}
f(u) - f(v) - \nabla f(v)^T (u - v) &= \sum_{i=1}^{n} u_i \log(u_i) - v_i \log(v_i) - (1 + \log(v_i))(u_i - v_i) \\
&= \sum_{i=1}^{n} u_i \left[\log(u_i) - \log(v_i)\right] + v_i - u_i \\
&= \sum_{i=1}^{n} u_i \log\left(\frac{u_i}{v_i}\right) + v_i - u_i \\
&= D_{KL}(u,v)
\end{aligned}$$

- Since f is strictly convex, then, $\forall\ u, v \in \mathbb{R}^n_{++}$ such that $\boxed{u \neq v}$,

  $$f(u) > f(v) + \nabla f(v)^T (u - v) \qquad \text{(1}^{\text{st}}\text{ order condition)}$$

  ie $\quad D_{KL}(u,v) > 0$.

Moreover, $\forall u \in \mathbb{R}^n_{++}$: $D_{KL}(u,u) = 0$.

Thus, $$\boxed{\begin{array}{l} D_{KL}(u,v) \geq 0 \quad \forall\ u,v \in \mathbb{R}^n_{++} \\ \text{and } D_{KL}(u,v) = 0 \iff u = v \end{array}}$$

Exercice n°3:

. Since $A \in S_n^+$, then, $\exists\, V \in S_n$ such that $A = V^T V$.

In fact, using the spectral theorem, $\exists\, U \in O_n$ and D a diagonal matrix such that: $A = U^T D U$ where $D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$
$0 \leq \lambda_1 \leq \dots \leq \lambda_n$ are the eigenvalues of A.

Let $V := U^T \mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})\, U \in S_n$.

$\Rightarrow V^2 = V^T V = U^T D U = A$.

• Let us reformulate the given problem to a SOCP:

$$(\mathcal{P}): \begin{cases} \min\; c^T x \\ \text{s.t}\;\; x^T(V^T V - bb^T)x \leq 0 \\ \quad\; b^T x \geq 0 \\ \quad\; Dx = g \end{cases} \iff \begin{cases} \min\; c^T x \\ \text{s.t}\;\; \|Vx\|_2^2 \leq (b^T x)^2 \\ \quad\; b^T x \geq 0 \\ \quad\; Dx = g \end{cases}$$

$$\iff \begin{cases} \min\; c^T x \\ \text{s.t}\;\; \|Vx\|_2 \leq |b^T x| \\ \quad\; b^T x \geq 0 \\ \quad\; Dx = g \end{cases}$$

$$\iff \begin{cases} \min\; c^T x \\ \text{s.t}\;\; \|Vx\|_2 \leq b^T x \\ \quad\; Dx = g \end{cases}$$

So, $(\mathcal{P})$ is a SOCP, then, $\boxed{(\mathcal{P}) \text{ is a convex optimization problem}}$.

• $$(\mathcal{P}) \iff \begin{cases} \min\limits_{x,z}\; c^T x \\ \text{s.t}\;\; \|z\|_2 \leq b^T x \\ \quad\; z = Vx \\ \quad\; Dx = g \end{cases}$$

* We derive the Lagrangian:

Let $x \in \mathbb{R}^n$, $z \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$, $\nu \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^n$.

$$\mathcal{L}(x, z, \alpha, \nu, \lambda) = c^T x + \alpha(\|z\|_2 - b^T x) + \lambda^T(Vx - z) + \nu^T(Dx - g)$$
$$= \left[c + V\lambda + D^T\nu - \alpha b\right]^T x + \left(\alpha\|z\|_2 - \lambda^T z\right) - \nu^T g$$

$$g(\alpha, \nu, \lambda) = \inf_{x,z} \mathcal{L}(x, z, \alpha, \nu, \lambda)$$
$$= \inf_x \left[c + V\lambda + D^T\nu - \alpha b\right]^T x + \inf_z \left(\alpha \|z\|_2 - \lambda^T z\right) - \nu^T g$$

- $\inf_x \left[c + V\lambda + D^T\nu - \alpha b\right]^T x = \begin{cases} 0 & \text{if } \alpha b = c + V\lambda + D^T\nu \\ -\infty & \text{otherwise} \end{cases}$

- Let $\varphi : z \in \mathbb{R}^m \longmapsto \alpha \|z\|_2 - \lambda^T z$.
  - If $\|\lambda\|_2 > \alpha$: $\forall t > 0$, $\varphi(t\lambda) = t\|\lambda\|_2(\alpha - \|\lambda\|_2) \xrightarrow[t \to +\infty]{} -\infty$: unbounded.
  - If $\|\lambda\|_2 \leqslant \alpha$: With Cauchy-Schwartz inequality, we have:
$$\lambda^T z \leqslant \|\lambda\|_2 \|z\|_2 \leqslant \alpha \|z\|_2$$
ie $\varphi(z) \geqslant 0$.

    Moreover $\varphi(0) = 0$.

  Then, $\inf_z \varphi(z) = \begin{cases} 0 & \text{if } \alpha \geqslant \|\lambda\|_2 \\ -\infty & \text{otherwise.} \end{cases}$

Thus, $g(\alpha, \nu, \lambda) = \begin{cases} -\nu^T g & \text{if } \alpha \geqslant \|\lambda\|_2 \text{ and } \alpha b = c + V\lambda + D^T\nu \\ -\infty & \text{otherwise.} \end{cases}$

$\Rightarrow$ We can derive the following dual problem of $(\mathcal{P})$:

$$\begin{cases} \max & g(\alpha, \nu, \lambda) \\ \text{s.t} & \alpha \geqslant 0 \end{cases} \iff \begin{cases} \max & -\nu^T g \\ \text{s.t.} & \alpha \geqslant 0 \\ & \alpha \geqslant \|\lambda\|_2 \\ & \alpha b = c + V\lambda + D^T\nu \end{cases}$$

$$\iff \boxed{\begin{cases} \min & g^T\nu \\ \text{s.t} & c + V\lambda + D^T\nu \succcurlyeq \|\lambda\|_2 b \end{cases}}$$

Exercice n°4 :

$$(\mathcal{P}) : \begin{cases} \min & -\sum_{i=1}^{m} \log(b_i - a_i^T x) \\ \text{s.t.} & a_i^T x - b_i < 0 \quad \forall i \in [\![1, m]\!] \end{cases}$$

Let $A := \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \in \mathbb{R}^{m \times n}$

$$(\mathcal{P}) \iff \begin{cases} \min\limits_{\substack{x \in \mathbb{R}^n \\ y \in \mathbb{R}^m_{++}}} & -\sum_{i=1}^{m} \log(y_i) \\ \text{s.t.} & y = b - Ax \end{cases}$$

□ The Lagrangian :

Let $x \in \mathbb{R}^n$, $\nu \in \mathbb{R}^m$, $y \in \mathbb{R}^m_{++}$:

$$\mathcal{L}(x, y, \nu) = -\sum_{i=1}^{m} \log(y_i) + \nu^T (y + Ax - b)$$

$$= -\sum_{i=1}^{m} \log(y_i) + \nu^T y + (A^T \nu)^T x - b^T \nu .$$

• Let $\varphi : y \in \mathbb{R}^m_{++} \mapsto \sum_{i=1}^{m} \log(y_i) - \nu^T y$.

$$g(\nu) = \inf_{\substack{x \in \mathbb{R}^n \\ y \in \mathbb{R}^m_{++}}} \mathcal{L}(x, y, \nu) = -\sup_{y \in \mathbb{R}^m_{++}} \varphi(y) + \inf_{x \in \mathbb{R}^n} \left((A^T\nu)^T x - b^T \nu \right)$$

• $$\inf_{x \in \mathbb{R}^n} (A^T \nu)^T x - b^T \nu = \begin{cases} -b^T \nu & \text{if } A^T \nu = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

• ▶ if $\exists i \in [\![1, m]\!]$ such that $\nu_i \leq 0$:

Then, $\forall t > 0$, $\varphi(t e_i) = \log(t) - \nu_i t \xrightarrow[t \to +\infty]{} +\infty$: unbounded

(where $(e_i)_{1 \leq i \leq m}$ is the canonical basis of $\mathbb{R}^m$)

▶ Let $\nu > 0$:

. $\forall i \in [\![1, m]\!]$: $\dfrac{\partial \varphi(y)}{\partial y_i} = \dfrac{1}{y_i} - \nu_i$; $\dfrac{\partial^2 \varphi(y)}{\partial y_i \partial y_j} = \begin{cases} -\dfrac{1}{y_i^2} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$

ie $\nabla^2 \varphi(y) = \operatorname{diag}\left(-\dfrac{1}{y_1^2}, \dots, -\dfrac{1}{y_m^2}\right) \prec 0$. Then, φ is strictly convex over $\mathbb{R}^m_{++}$.

.. $\nabla \varphi(y) = 0 \iff \forall i \in [\![1, m]\!]$, $y_i = \dfrac{1}{\nu_i}$

$\Rightarrow$ Then, $\sup_{y \in \mathbb{R}^m_{++}} \ell(y) = \begin{cases} \sum_{i=1}^{m} \log\left(\frac{1}{\nu_i}\right) - m & \text{if } \nu \succ 0 \\ +\infty & \text{otherwise.} \end{cases}$

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

A dual of $(\mathcal{P})$ is given by:

$$(D): \boxed{\begin{cases} \max_{\nu \in \mathbb{R}^m_{++}} & m + \sum_{i=1}^{m} \log(\nu_i) - b^T\nu \\ \text{s.t.} & A^T\nu = 0 \end{cases}}$$

Exercice n°5:

$(P_1): \begin{cases} \min f_0(x) \\ \text{s.t. } Ax = b \end{cases}$ $\qquad (P_2): \min f_0(x) + \alpha \|Ax - b\|_2^2$

- The Lagrangian associated to (P_1) is: $\mathcal{L}(x,\nu) = f_0(x) + \nu^T(Ax - b)$

  Then, $\nabla_x \mathcal{L}(x,\nu) = \nabla_x f_0(x) + A^T\nu$.

- If $\tilde{x}$ minimizes $\Phi : x \mapsto f_0(x) + \alpha\|Ax - b\|_2^2$, then, since Φ is differentiable, we have: $\nabla_x \Phi(\tilde{x}) = 0$

  ie $\nabla_x f_0(\tilde{x}) + 2\alpha A^T(A\tilde{x} - b) = 0$.

  ie $\tilde{x}$ satisfies $\nabla_x \mathcal{L}(\tilde{x}, 2\alpha(A\tilde{x} - b)) = 0$.

If we introduce $\boxed{\nu_0 := 2\alpha(A\tilde{x} - b)}$, then $\tilde{x}$ satisfies: $\nabla_x \mathcal{L}(\tilde{x}, \nu_0) = 0$

Since f_0 is convex, then, $x \mapsto \mathcal{L}(x, \nu_0)$ is convex (as a sum of two convex functions: f_0 and an affine transformation)

Therefore, $\tilde{x}$ is a minimizer of $x \mapsto \mathcal{L}(x, \nu_0)$

We conclude that $\boxed{\nu_0 \text{ is dual feasible}}$ with: $g(\nu_0) = \inf_x \mathcal{L}(x, \nu_0)$

$= \mathcal{L}(\tilde{x}, \nu_0)$

$g(\nu_0) = f_0(\tilde{x}) + 2\alpha\|A\tilde{x} - b\|_2^2$.

Using weak duality, for each x primal feasible (P_1) and each ν dual feasible, we have: $f_0(x) \geq g(\nu)$.

In particular, $f_0(x) \geq g(\nu_0)$ $\quad \forall x$ primal feasible

ie $\boxed{f_0(x) \geq f_0(\tilde{x}) + 2\alpha \|A\tilde{x} - b\|_2^2 \quad \forall x \in \mathbb{R}^n \text{ such that } Ax = b}$

Exercice n°6:

Let $X := \begin{pmatrix} x_1^T \\ \vdots \\ x_m^T \end{pmatrix} \in \mathbb{R}^{m \times n}$ and $Y = \operatorname{diag}(y_1, \dots, y_m) \in \mathbb{R}^{m \times m}$

Then, $U := YX = \begin{pmatrix} y_1 x_1^T \\ \vdots \\ y_m x_m^T \end{pmatrix}$

- The classification problem can be written as following:

$$\begin{cases} \min & \frac{1}{2}\|w\|_2^2 + C\, \mathbb{1}_m^T z \\ \text{st} & z \geq 0 \\ & \mathbb{1}_m - z \leq Uw \end{cases}$$

1

- The Lagrangian:

Let $w \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, $\lambda, \nu \in \mathbb{R}^m$:

$$\mathcal{L}(w, z, \lambda, \nu) = \frac{1}{2}\|w\|_2^2 + C\,\mathbb{1}_m^T z - \lambda^T z + \nu^T(\mathbb{1}_m - z - Uw)$$
$$= \frac{1}{2}\|w\|_2^2 - (U^T\nu)^T w + (C\,\mathbb{1}_m - \lambda - \nu)^T z + \mathbb{1}_m^T \nu$$

$$\Rightarrow g(\lambda, \nu) = \inf_{w,z} \mathcal{L}(w, z, \lambda, \nu)$$
$$= \inf_w \left(\frac{1}{2}\|w\|_2^2 - (U^T\nu)^T w\right) + \inf_z (C\,\mathbb{1}_m - \lambda - \nu)^T z + \mathbb{1}_m^T \nu.$$

- $\inf_z (C\,\mathbb{1}_m - \lambda - \nu)^T z = \begin{cases} 0 & \text{if } C\,\mathbb{1}_m = \lambda + \nu \\ -\infty & \text{otherwise.} \end{cases}$

- Let $\varphi : w \in \mathbb{R}^n \longmapsto \frac{1}{2}\|w\|_2^2 - (U^T\nu)^T w$

. φ is twice differentiable. $\quad \nabla\varphi(w) = w - U^T\nu$

$\nabla\varphi(w) = 0 \iff w = U^T\nu$

. $\nabla^2\varphi(w) = I_n \succ 0$. Then, φ is convex.

Thus, $\inf_w \varphi(w) = \varphi(U^T\nu) = -\frac{1}{2}\|U^T\nu\|^2$

Therefore, $g(\lambda,\nu) = \begin{cases} -\frac{1}{2}\|W^T\nu\|^2 + \mathbb{1}_m^T\nu & \text{if } C\mathbb{1}_m = \lambda+\nu \\ -\infty & \text{otherwise} \end{cases}$

Then, the dual is given by:

$$\begin{cases} \max & g(\lambda,\nu) \\ \text{s.t} & \lambda \geqslant 0 \\ & \nu \geqslant 0 \end{cases} \iff \begin{cases} \max & -\frac{1}{2}\|W^T\nu\|^2 + \mathbb{1}_m^T\nu \\ \text{s.t} & \lambda \geqslant 0 \\ & \nu \geqslant 0 \\ & C\mathbb{1}_m = \lambda+\nu \end{cases}$$

$$\iff \boxed{\begin{cases} \min_\nu & \frac{1}{2}\|W^T\nu\|^2 - \mathbb{1}_m^T\nu \\ \text{s.t} & 0 \leqslant \nu \leqslant C\mathbb{1}_m \end{cases}}$$

2

It is not obvious if the question requires to apply separately the barrier method to the primal and the dual or to derive the barrier method to only one of them and then deduce the solution of the other problem by the KKT optimality conditions. The second approach is clearly the easiest one especially if we apply the barrier method to the dual: in fact, it is easier to derive the centering step of the barrier method associated to the dual problem than the one corresponding to the primal.

Barrier Method Algorithm

- Starting point: ν_0 strictly feasible (for example: $\lambda_0 = \frac{C}{2}\mathbb{1}_m$); $t_0 > 0$ (for exp: $t_0 = \frac{m}{\varepsilon}$)
  $\mu > 0$; tolerance: $\varepsilon > 0$

- Repeat: 1 Centering step: compute $\nu^*(t) := \arg\min_\nu \underbrace{t f_0(\lambda) + \Phi(\lambda)}_{\text{we will note it } F_t(\lambda)}$
  (by Newton method)

  where $f_0 : \nu \longmapsto \frac{1}{2}\|W^T\nu\|^2 - \mathbb{1}_m^T\nu$

  and $\Phi : \nu \longmapsto -\sum_{i=1}^m \log(\nu_i) - \sum_{i=1}^m \log(C-\nu_i)$

  a) Compute: • normalized Newton step: $\Delta\nu_{nt} = -\nabla^2 F_t(\lambda)\,\nabla F_t(\lambda)$

  • Newton decrement: $\lambda^2 = -\nabla F_t(\lambda)^T \Delta\nu_{nt}$

  b) Stopping criterion: quit if $\lambda^2 \leqslant 2\varepsilon$

  c) Line Search: choose h by backtracking (hyperparameters: α, β)

d) update : $\nu \leftarrow \nu + h \times \Delta\nu_{nt}$

[2] update : $\nu \leftarrow \nu^*(t)$

[3] Stopping criterion : quit if $t > \frac{m}{\varepsilon}$

[4] update : $t \leftarrow \mu t$

* Details of the centering step : the goal of the centering step is to solve the unconstrained problem : $\min_{\nu} F_t(\nu)$

where $F_t(\nu) = t\left(\frac{1}{2}\|U^T\nu\|^2 - \mathbb{1}_m^T \nu\right) - \sum_{i=1}^{m} \log(\nu_i) - \sum_{i=1}^{m} \log(c - \nu_i)$

- $\boxed{\nabla F_t = t\left(UU^T\nu - \mathbb{1}_m\right) - \frac{1}{\nu} + \frac{1}{c-\nu}}$
- $\boxed{\nabla^2 F_t = t\,UU^T + \operatorname{diag}\left(\frac{1}{\nu_1^2} + \frac{1}{(c-\nu_1)^2}, \dots, \frac{1}{\nu_m^2} + \frac{1}{(c-\nu_m)^2}\right)}$

Notation : $\frac{1}{\nu} := \left[\frac{1}{\nu_1}, \dots, \frac{1}{\nu_m}\right]^T \in \mathbb{R}^m$

- Remark about the backtrack line search inside the centering criterion :
  - Starting point : h_0 such that $\nu + h\,\Delta\nu_{nt} \in \operatorname{dom} F_t$
  - Repeat : $h \leftarrow \beta \times h$

    Until : $(\nu + h\,\Delta\nu_{nt}) \notin \operatorname{dom} F_t$ or $F_t(\nu + h\,\Delta\nu_{nt}) < F_t(\nu) + \alpha h \nabla F_t(\nu)^T \Delta\nu_{nt}$

We choose h_0 such that $0 < \nu + h_0\,\Delta\nu_{nt} < c$ ie $\begin{cases} \text{if } [\Delta\nu_{nt}]_i > 0 : \frac{-\nu}{[\Delta\nu]_i} < h_0 < \frac{c - \nu_i}{[\Delta\nu]_i} \\ \text{if } [\Delta\nu]_i < 0 : \frac{c-\nu}{[\Delta\nu]_i} < h_0 < \frac{-\nu_i}{[\Delta\nu]_i} \end{cases}$

Since the stepsize h decreases over iterations, we can set :

$$h_0 = 0{,}99 \times \min\left(\min\left(\frac{c-\nu_i}{[\Delta\nu_{nt}]_i} \text{ for } [\Delta\nu_{nt}]_i > 0\right), \min\left(-\frac{\nu_i}{[\Delta\nu_{nt}]_i} \text{ for } [\Delta\nu]_i < 0\right)\right)$$

- The described barrier algorith yields an approximative solution $\boxed{\hat{\nu}}$ of the dual problem with a tolerance ε.

We know that the primal is strictly feasible (for example $\lambda_0 = \frac{c}{2}\mathbb{1}_m$), then, it satisfies the Slater's condition. Moreover, it is convex. Then, strong duality holds. Therefore, if $w^*, z^*, \nu^*, \lambda^*$ are optimal, then, they must satisfy the KKT optimality conditions :

1) Primal feasibility : $z^* \succcurlyeq 0$ and $\mathbb{1}_m \preccurlyeq Uw^* + z^*$

2) Dual feasibility : $\nu^* \succcurlyeq 0$ and $\lambda^* \succcurlyeq 0$

3) Complementary slackness : $\nu_i^*\left(1 - z_i^* - [Uw^*]_i\right) = 0$ and $\lambda_i^* z_i^* = 0$

4) First order condition : $w^* = U^T\nu^*$ and $c\,\mathbb{1}_m = \lambda^* + \nu^*$

Then, $\boxed{w^* = U^T\nu^*}$ and $\boxed{z_i^* = \begin{cases} 0 & \text{if } \nu_i^* < c \\ 1 - [Uw^*]_i & \text{if } \nu_i^* = c \end{cases}}$

Let $\hat{w} := U^T \hat{\nu}$ and $\hat{z} \in \mathbb{R}^m$ such that $\hat{z}_i = \begin{cases} 1 - [U\hat{w}]_i & \text{if } \hat{\nu}_i = c \\ 0 & \text{otherwise} \end{cases}$

- We notice that: $c \mathbb{1}_m^T \hat{z} = \hat{\nu}^T \hat{z} = \hat{\nu}^T \mathbb{1}_m - \hat{\nu}^T (U\hat{w})$
- Then, $f(\hat{w}, \hat{z}) - f(w^*, z^*) = f(\hat{w}, \hat{z}) - g(\nu^*)$ (because we have strong duality)

$$= \frac{1}{2} \|\hat{w}\|^2 + c \mathbb{1}_m^T \hat{z} - g(\nu^*)$$

$$= \frac{1}{2} \|U^T \hat{\nu}\|^2 + \mathbb{1}_m^T \hat{\nu} - \hat{\nu}^T U \hat{w} - g(\nu^*)$$

$$= \frac{1}{2} \|U^T \hat{\nu}\|^2 + \mathbb{1}_m^T \hat{\nu} - \hat{\nu}^T U (U^T \hat{\nu}) - g(\nu^*)$$

$$= -\frac{1}{2} \|U^T \hat{\nu}\|^2 + \mathbb{1}_m^T \hat{\nu} - g(\nu^*)$$

$$= g(\hat{\nu}) - g(\nu^*).$$

By definition of $\hat{\nu}$, $g(\nu^*) - g(\hat{\nu}) \leqslant \varepsilon$. Thus, $f(w^*, z^*) - f(\hat{w}, \hat{z}) \leqslant \varepsilon$. which means that $(\hat{w}, \hat{z})$ is an approximative solution to the primal problem with tolerance ε.