MA1407.000500 Mathematical Algorithms II Assignment 1

Fall 2025

Out: 18th September 2025

Due: 1st October 2025 at 11:59pm

Instructions

- Submit a PDF document with your solutions through the assignment portal on KLMS by the due date. Please ensure that your name and student ID are on the front page.
- Late assignments will **not** be accepted except in extenuating circumstances. Special consideration should be applied for in this case.
- It is **required** that you typeset your solutions in LaTeX. Handwritten solutions will not be accepted.
- Spend some time ensuring your arguments are **coherent** and your solutions **clearly** communicate your ideas.

Question:	1	2	3	4	5	Total
Points:	15	30	15	20	20	100

1. This question asks you to prove that $f(x) = \|Ax - b\|_2^2/(1 - \|x\|_2^2)$ is convex on the open unit ball $X = \{x \in \mathbb{R}^d : \|x\|_2 < 1\}$. Answer the following sub-questions that lead to a proof.
   (a) (5 points) Show that $g(x,t) = \|x\|_2^2 + \|Ax - b\|_2^2/t$ over $(x,t) \in X \times \mathbb{R}_{++}$ is convex.
   (b) (5 points) We define the epigraph of f as $\text{epi}(f) = \{(x,t) : f(x) \le t\}$. Show that the epigraph of f is given by
   $$\text{epi}(f) = \{(x,t) : g(x,t) \le 1, \|x\|_2 < 1, t > 0\} \cup \{(x,0) : \|x\|_2 < 1, Ax = b\}.$$
   (c) (5 points) Complete the proof that f is convex on X.

Solution:

(a) The perspective of $\|x\|_2^2$ is $\|x\|_2^2/t$ is convex, therefore $\|Ax-b\|_2^2/t$ is convex since it is obtained from the perspective via a linear transformation. Clearly, $g(x,t) = \|x\|_2^2 + \|Ax - b\|_2^2/t$ is also convex.

(b) Now, for $t > 0$,
$$(x,t) \in \text{epi}(f) \iff \frac{\|Ax-b\|_2^2}{1 - \|x\|_2^2} \le t \iff \frac{\|Ax-b\|_2^2}{t} \le 1 - \|x\|_2^2.$$
Furthermore, for $t = 0$, $(x,t) \in \text{epi}(f)$ if and only if $Ax = b$ and $\|x\|_2^2 < 1$.
Therefore
$$\text{epi}(f) = \{(x,t) : g(x,t) \le 1, \|x\|_2 < 1, t > 0\} \cup \{(x,0) : \|x\|_2 < 1, Ax = b\}.$$

(c) Each component is clearly convex. Now consider (x,t) such that $g(x,t) \le 1$, $\|x\|_2 < 1$ and $t > 0$, as well as x' such that $\|x\|_2 < 1$, $Ax' = b$. Take $\lambda \in (0,1)$ and consider the convex combination
$$(\lambda x + (1-\lambda)x', \lambda t).$$
Note that
$$\begin{aligned}
g(\lambda x + (1-\lambda)x', \lambda t) &= \|\lambda x + (1-\lambda)x'\|_2^2 + \frac{\|A(\lambda x + (1-\lambda)x') - b\|_2^2}{\lambda t} \\
&= \|\lambda x + (1-\lambda)x'\|_2^2 + \frac{\|\lambda(Ax - b) + (1-\lambda)(Ax' - b)\|_2^2}{\lambda t} \\
&= \|\lambda x + (1-\lambda)x'\|_2^2 + \frac{\lambda\|Ax-b\|_2^2}{t} \\
&\le \lambda\|x\|_2^2 + (1-\lambda)\|x'\|_2^2 + \frac{\lambda\|Ax-b\|_2^2}{t} \\
&= (1-\lambda)\|x'\|_2^2 + \lambda g(x,t) \\
&\le 1.
\end{aligned}$$
We also clearly have $\lambda t > 0$ and $\|\lambda x + (1-\lambda)x'\|_2^2 < 1$. Therefore
$$(\lambda x + (1-\lambda)x', \lambda t) \in \text{epi}(f).$$
We conclude that epi(f) is convex.

2. Verify convexity/concavity of the following functions. You may use the first-order and second-order characterizations of convex functions, while there exists a direct proof based on the definition of convex functions.
   (a) (5 points) The *negative entropy function* is convex on $\mathbb{R}_{++}^d$:
   $$f(x) := \sum_{i\in[d]} x_i \log(x_i).$$

Solution: It suffices to show that $g(z) = z\log(z)$ is convex on $\mathbb{R}_{++}$. Note that
$$\max_y \{yz - \exp(y)\} = z(\log(z) - 1).$$

Therefore $z\log(z) = \max_y \{yz - \exp(y)\} + z$ is convex.

(b) (5 points) The *log-sum-exp function* is convex:

$$f(x) = \log\left(\sum_{i\in[d]} \exp(x_i)\right).$$

[Hint: An elementary proof exists by showing $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$. You may use (without proof) the inequality $\sum_{i\in[d]} |u_i|^\lambda |v_i|^{1-\lambda} \le \left(\sum_{i\in[d]} |u_i|\right)^\lambda \left(\sum_{i\in[d]} |v_i|\right)^{1-\lambda}$.]

Solution: Consider points x, y and weight $\lambda \in [0,1]$. Then

$$\begin{aligned}
\log\left(\sum_{i\in[d]} \exp\left(\lambda x_i + (1-\lambda)y_i\right)\right) &= \log\left(\sum_{i\in[d]} \exp\left(x_i\right)^\lambda \exp\left(y_i\right)^{1-\lambda}\right) \\
&\le \log\left(\left(\sum_{i\in[d]} \exp\left(x_i\right)\right)^\lambda \left(\sum_{i\in[d]} \exp\left(y_i\right)\right)^{1-\lambda}\right) \\
&= \lambda \log\left(\sum_{i\in[d]} \exp\left(x_i\right)\right) + (1-\lambda)\log\left(\sum_{i\in[d]} \exp\left(x_i\right)\right).
\end{aligned}$$

(c) (5 points) The geometric mean is concave on $\mathbb{R}^d_{++}$:

$$f(x) = \left(\prod_{i\in[d]} x_i\right)^{1/d}.$$

Solution: Given $x, y \in \mathbb{R}^d_{++}$ and $\lambda \in [0,1]$ we have

$$\begin{aligned}
\frac{\lambda f(x) + (1-\lambda)f(y)}{f(\lambda x + (1-\lambda)y)} &= \lambda\left(\prod_{i\in[d]} \frac{x_i}{\lambda x_i + (1-\lambda)y_i}\right)^{1/d} + (1-\lambda)\left(\prod_{i\in[d]} \frac{y_i}{\lambda x_i + (1-\lambda)y_i}\right)^{1/d} \\
&\le \frac{\lambda}{d}\sum_{i\in[d]} \frac{x_i}{\lambda x_i + (1-\lambda)y_i} + \frac{1-\lambda}{d}\sum_{i\in[d]} \frac{y_i}{\lambda x_i + (1-\lambda)y_i} \\
&= 1
\end{aligned}$$

where the inequality follows from the arithmetic-geometric mean inequality. Therefore f is concave.

(d) (5 points) The log-determinant is concave on $\mathbb{S}^d_{++}$, the set of $d \times d$ positive "definite" matrices:

$$f(X) = \log\det(X).$$

Note: the determinant of a matrix $X \in \mathbb{S}^d_{++}$ is simply the product of its eigenvalues (which are all positive by assumption). You will need the following properties of matrices/determinants which you can use without proof:

- $\det(AB) = \det(BA) = \det(A)\det(B)$ for $A, B \in \mathbb{S}^d_{++}$.
- If $A \in \mathbb{S}^d_{++}$, then we can write $A = PDP^\top$ where $P^\top P = PP^\top = I$ and D is diagonal with strictly positive entries. Let D^r be the diagonal matrix with all diagonal entries of D raised to the power $r \in \mathbb{R}$. We can define powers of A via $A^r = PD^rP^\top$, which have the same properties as the usual powers: $A^rA^s = A^sA^r = A^{r+s}$, $A^0 = I$. Furthermore, $\det(A^r) = \det(A)^r$.
- $\alpha A + \beta B = A^{1/2}(\alpha I + \beta A^{-1/2}BA^{-1/2})A^{1/2}$ for any $A, B \in \mathbb{S}^d_{++}$ and $\alpha, \beta \ge 0$.

Solution: Consider $X, Y \in \mathbb{S}^d_{++}$ and $\lambda \in [0,1]$. Now

$$\begin{aligned}\log\det(\lambda X + (1-\lambda)Y) &= \log\det(Y + \lambda(X-Y)) \\ &= \log\det(Y^{1/2}(I + \lambda Y^{-1/2}(X-Y)Y^{-1/2})Y^{1/2}) \\ &= \log\det(I + \lambda Y^{-1/2}(X-Y)Y^{-1/2}) + \log\det(Y) \\ &= \log\det((1-\lambda)I + \lambda Y^{-1/2}XY^{-1/2}) + \log\det(Y).\end{aligned}$$

Note that the eigenvalues of $I + \lambda Y^{-1/2}(X-Y)Y^{-1/2}$ are $1-\lambda+\lambda\gamma_i$, where γ_i are eigenvalues of $Y^{-1/2}XY^{-1/2}$. Note that $Y^{-1/2}XY^{-1/2} \in \mathbb{S}^d_{++}$ also, therefore $1-\lambda+\lambda\gamma_i > 0$. Hence

$$\log\det(\lambda X + (1-\lambda)Y) = \log\det(Y) + \sum_{i\in[d]} \log(1-\lambda+\lambda\gamma_i).$$

When $\lambda = 0$, then this is clearly $\log\det(Y)$, and when $\lambda = 1$, we have

$$\begin{aligned}\log\det(Y) + \sum_{i\in[d]} \log(1-\lambda+\lambda\gamma_i) &= \log\det(Y) + \sum_{i\in[d]} \log(\gamma_i) \\ &= \log\det(Y) + \log\det(Y^{-1/2}XY^{-1/2}) \\ &= \log\det(X)\end{aligned}$$

where the last equality follows since $\log\det(Y^{-1/2}XY^{-1/2}) = \log\det(Y^{-1/2}) + \log\det(X) + \log\det(Y^{-1/2}) = \log\det(X) - \log\det(Y)$. As $\log\det(Y) + \sum_{i\in[d]} \log(1-\lambda+\lambda\gamma_i)$ is a concave function of λ, we have

$$\begin{aligned}\log\det(\lambda X + (1-\lambda)Y) &= \log\det(Y) + \sum_{i\in[d]} \log(1-\lambda+\lambda\gamma_i) \\ &\geq \lambda\left(\log\det(Y) + \sum_{i\in[d]} \log(\gamma_i)\right) + (1-\lambda)\log\det(Y) \\ &= \lambda\log\det(X) + (1-\lambda)\log\det(Y).\end{aligned}$$

(e) (5 points) The *conjugate* of a function $f : \mathbb{R}^d \to \mathbb{R}$:

$$f^*(x) = \sup_{y\in\mathbb{R}^d} \{\langle y, x\rangle - f(y)\}$$

Solution: For a fixed $y \in \mathbb{R}^d$, note that $\langle y, x\rangle - f(y)$ is a linear function and thus convex. Note that

$$\begin{aligned}\operatorname{epi}(f^*) &= \{(x,t) : \sup_{y\in\mathbb{R}^d} \{\langle y, x\rangle - f(y)\} \leq t\} \\ &= \bigcap_{y\in\mathbb{R}^d} \{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \langle y, x\rangle - f(y) \leq t\}.\end{aligned}$$

As the epigraph of $\langle y, x\rangle - f(y)$ is convex and $\operatorname{epi}(f^*)$ is the intersection of convex sets, $\operatorname{epi}(f^*)$ is convex. As a result, f^* is convex.

(f) (5 points) The sum of k largest components of $x \in \mathbb{R}^d$:

$$f(x) = x_{\sigma(1)} + \cdots + x_{\sigma(k)}$$

where $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(d)}$ are the rearrangement of $x_1, \ldots, x_d$ in nonincreasing order.

Solution: Take any combination of k distinct indices $1 \leq i_1 < i_2 < \cdots < i_k \leq d$. Then

$$x_{i_1} + \cdots + x_{i_k}$$

is clearly a convex function. Note that

$$f(x) = x_{\sigma(1)} + \cdots + x_{\sigma(k)} = \max_{i_1,\ldots,i_k} \{x_{i_1} + \cdots + x_{i_k} : 1 \leq i_1 < i_2 < \cdots < i_k \leq d\} .$$

Hence f is the point-wise maximum of convex functions, and therefore, f is convex.

3. Let $K = \{(x, y, z) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R} : y \exp(x/y) \leq z\}$.
   (a) (5 points) Let $K = \{(x, y, z) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R} : y \exp(x/y) \leq z\}$. Show that K is a convex cone.
   (b) (2 points) The dual cone of K is given by

   $$K^* = \left\{(u, v, w) \in \mathbb{R}^3 : ux + vy + wz \geq 0 \ \forall (x, y, z) \in K\right\} .$$

   Show that for any $(u, v, w) \in K^*$ with $w = 0$, the point satisfies $u = 0$ and $v \geq 0$. Moreover, argue that $\{(0, v, 0) : v \geq 0\} \subset K^*$.
   (c) (2 points) Show that for any $(u, v, w) \in K^*$ with $w > 0$ and $u = 0$, the point satisfies $v \geq 0$. Moreover, argue that $\{(0, v, 0) : v \geq 0\} \subset K^*$. Moreover, argue that $\{(0, v, w) : v \geq 0, w > 0\} \subset K^*$.
   (d) (2 points) Show that for any $(u, v, w) \in K^*$ with $w > 0$ and $u \neq 0$, the point satisfies $u < 0$.
   (e) (4 points) Finish the proof that the dual cone of K from part (b) is given by

   $$K^* = \{(u, v, w) : u < 0, -u \exp(v/u) \leq \exp(1)w\} \cup \{(0, v, w) : v, w \geq 0\} .$$

   [Hint: argue that if $w > 0$ and $u < 0$, then $-u \exp v/u \leq \exp(1)w$ holds. You can use the fact (without proof) that minimizers of a differentiable convex function $f : \mathbb{R} \to \mathbb{R}$ occur at points where $f'(x) = 0$. We prove a more general version of this in lectures.]

Solution:

(a) Note that $y \exp(x/y)$ is the perspective of the exponential function $\exp(\cdot)$, hence is convex. Then K is simply the epigraph of the perspective, so it is also convex. To show that it is a cone, observe that if $y \exp(x/y) \leq z$, then $(\alpha y) \exp((\alpha x)/(\alpha y)) = \alpha y \exp(x/y) \leq \alpha z$ for $\alpha > 0$.

(b) The dual cone is

$$K^* = \left\{(u, v, w) \in \mathbb{R}^3 : ux + vy + wz \geq 0 \ \forall (x, y, z) \in K\right\} .$$

Since we can make z arbitrarily large in K, we must have $w \geq 0$.

When $w = 0$, we can make y arbitrarily large as well, so we must have $v \geq 0$. When $v = 0$, we require $ux \geq 0$, and since x can be positive or negative, we must have $u = 0$. Therefore $(0, 0, 0) \in K^*$. When $v > 0$, we have $ux + vy \geq 0$. We can fix y and make x arbitrarily positive or negative, therefore we need $u = 0$ again. Thus, $\{(0, v, 0) : v \geq 0\} \subset K^*$.

(c) Now consider $w > 0$. We require $ux + vy + wz \geq 0$ for all $(x, y, z) \in K$. Without loss of generality, we can take $z = y \exp(x/y)$, thus we require

$$ux + vy + wy \exp(x/y) \geq 0 \quad \forall x \in \mathbb{R}, y > 0.$$

Fix y for the moment. If $u = 0$, then by making x arbitrarily small, we deduce that we require $v \geq 0$, i.e., $\{(0, v, w) : v \geq 0, w > 0\} \subset K^*$.

(d) If $u > 0$, then we can make x arbitrarily small to make the term arbitrarily small. Therefore we can't have $u > 0$.

(e) Now we consider $u < 0$. By differentiating and setting to 0, the minimum x occurs at $\exp(x/y) = -u/w$, thus substituting this in we get

$$y(u \log(-u/w) + v - u) \geq 0.$$

Therefore we require $u \log(-u/w) + v - u \geq 0$. Simplifying this, we get $-u \exp(v/u) \leq \exp(1)w$.

Therefore we have

$$K^* = \{(u, v, w) : u < 0, -u \exp(v/u) \leq \exp(1)w\} \cup \{(0, v, w) : v, w \geq 0\} .$$

4. Consider the following algorithm.

Algorithm 1 AdaGrad

Initialize $x_1 \in \mathcal{X}$ and $S_0 = 0$.
for $t = 1, \dots, T$ **do**
  Obtain a subgradient $g_t \in \partial f(x_t)$.
  Set $S_t = S_{t-1} + \|g_t\|_2^2$ and $\eta_t = R/\sqrt{2S_t}$.
  Update $x_{t+1} = \mathrm{Proj}_{\mathcal{X}}(x_t - \eta_t g_t)$.
end for

(a) (5 points) Let $\{x_t : t = 1, \dots, T\}$ be the sequence of iterates generated by AdaGrad. Then

$$\sum_{t=1}^{T} g_t^\top (x_t - x^*) \leq \sqrt{2R^2 \sum_{t=1}^{T} \|g_t\|_2^2}.$$

Solution: Note that

$$\begin{aligned} \|x_{t+1} - x^*\|_2^2 &\leq \|x_t - \eta_t g_t - x^*\|_2^2 \\ &= \|x_t - x^*\|_2^2 - 2\eta_t g_t^\top (x_t - x^*) + \eta_t^2 \|g_t\|_2^2. \end{aligned}$$

Then it follows that

$$\begin{aligned} \sum_{t=1}^{T} g_t^\top (x_t - x^*) &\leq \frac{\|x_1 - x^*\|_2^2}{2\eta_1} + \sum_{t=1}^{T} \frac{\|x_t - x^*\|_2^2}{2} \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) + \sum_{t=1}^{T} \frac{\eta_t}{2} \|g_t\|_2^2 \\ &\leq \frac{R^2}{2\eta_T} + \sum_{t=1}^{T} \frac{\eta_t}{2} \|g_t\|_2^2 \\ &\leq \frac{R}{2}\sqrt{2S_T} + \frac{R}{2\sqrt{2}} \sum_{t=1}^{T} \frac{\|g_t\|_2^2}{\sqrt{S_t}}. \end{aligned}$$

It is known that for any non-negative numbers $a_1, \dots, a_n$, we have

$$\sum_{i=1}^{n} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} \leq 2\sqrt{\sum_{i=1}^{n} a_i}.$$

Applying this inequality, it follows that

$$\sum_{t=1}^{T} g_t^\top (x_t - x^*) \leq \frac{R}{2}\sqrt{2S_T} + \frac{R}{\sqrt{2}}\sqrt{S_T} = \sqrt{2R^2 \sum_{t=1}^{T} \|g_t\|_2^2},$$

as required.

(b) (5 points) Suppose that f is convex with $\|g\|_2 \leq L$ for any $g \in \partial f(x)$ for every $x \in \mathbb{R}^d$. Let $\{x_t : t = 1, \dots, T\}$ be the sequence of iterates generated by AdaGrad. Then

$$f\left(\frac{1}{T} \sum_{t=1}^{T} x_t \right) - f(x^*) \leq \frac{LR\sqrt{2}}{\sqrt{T}}.$$

Solution: Note that

$$f\left(\frac{1}{T} \sum_{t=1}^{T} x_t \right) - f(x^*) \leq \frac{1}{T} \left(\sum_{t=1}^{T} (f(x_t) - f(x^*)) \right) \leq \frac{1}{T} \left(\sum_{t=1}^{T} g_t^\top (x_t - x^*) \right)$$

due to convexity of f. By part (a), we have

$$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \le \frac{1}{T}\sqrt{2R^2\sum_{t=1}^{T}\|g_t\|_2^2} \le \frac{LR\sqrt{2}}{\sqrt{T}},$$

as required.

(c) (5 points) Suppose that f is convex and β-smooth in the ℓ_2-norm. Then

$$f(y) \ge f(x) + \nabla f(x)^\top (y-x) + \frac{1}{2\beta}\|\nabla f(x) - \nabla f(y)\|_2^2.$$

Solution: Since f is β-smooth, we have

$$f(z) \le f(x) + \nabla f(x)^\top (z-x) + \frac{\beta}{2}\|z-x\|_2^2 \quad \text{for any } z \in \mathbb{R}^d.$$

Then taking $z = y + (1/\beta)(\nabla f(x) - \nabla f(y))$,

$$\begin{aligned} f(x) - f(y) &= f(x) - f(z) + f(z) - f(y) \\ &\le -\nabla f(x)^\top (z-x) + \nabla f(y)^\top (z-y) + \frac{\beta}{2}\|z-y\|_2^2 \\ &= \nabla f(x)^\top (x-y) + (\nabla f(x) - \nabla f(y))^\top (y-z) + \frac{\beta}{2}\|z-y\|_2^2 \\ &= \nabla f(x)^\top (x-y) - \frac{1}{2\beta}\|\nabla f(x) - \nabla f(y)\|_2^2. \end{aligned}$$

Then it follows that

$$\frac{1}{2\beta}\|\nabla f(x) - \nabla f(y)\|_2^2 \le f(y) - f(x) + \nabla f(x)^\top (x-y),$$

as required.

(d) (5 points) Suppose that f is convex and β-smooth in the ℓ_2-norm. Moreover, assume that $\sup_{x,y\in\mathcal{X}} \|x-y\|_2 \le R$. Let $\{x_t : t = 1, \ldots, T\}$ be the sequence of iterates generated by AdaGrad. Then

$$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \le \frac{\beta R^2}{T}$$

where x^* is an optimal solution to $\min_{x\in\mathbb{R}^d} f(x)$.

Solution: We have

$$f(x_t) - f(x^*) \le \nabla f(x_t)^\top (x_t - x^*) - \frac{1}{2\beta}\|\nabla f(x_t)\|_2^2,$$

which implies that

$$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \le \frac{1}{T}\sum_{t=1}^{T} \nabla f(x_t)^\top (x_t - x^*) - \frac{1}{2\beta T}\sum_{t=1}^{T}\|\nabla f(x_t)\|_2^2.$$

Moreover, by part(c), we deduce that

$$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \le \frac{1}{T}\sqrt{2R^2\sum_{t=1}^{T}\|\nabla f(x_t)\|_2^2} - \frac{1}{2\beta T}\sum_{t=1}^{T}\|\nabla f(x_t)\|_2^2.$$

Here,

$$\sqrt{2R^2\sum_{t=1}^{T}\|\nabla f(x_t)\|_2^2} - \frac{1}{2\beta}\sum_{t=1}^{T}\|\nabla f(x_t)\|_2^2 \le \max_{z\ge 0}\left\{Rz\sqrt{2} - \frac{1}{2\beta}z^2\right\} = \beta R^2.$$

Therefore, it follows that

$$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \leq \frac{\beta R^2}{T},$$

as required.

5. (20 points) Consider the ℓ_0-norm $\|x\|_0$ which is defined as the number of non-zero entries in $x \in \mathbb{R}^n$. Derive the proximal operator of the ℓ_0-norm with parameter $\eta > 0$:

$$\mathrm{prox}_{\eta\|\cdot\|_0}(x) = \arg\min_{y\in\mathbb{R}^n} \left\{\frac{1}{2}\|u - x\|_2^2 + \eta\|u\|_0\right\}.$$

Solution: Both terms $\|u - x\|_2^2 = \sum_{i=1}^n (u_i - x_i)^2$ and $\|u\|_0 = \sum_{i=1}^n \mathbf{1}_{\{u_i \neq 0\}}$ are separable across coordinates. Hence the problem decouples into n scalar problems:

$$u_i \in \arg\min_{t\in\mathbb{R}} \left\{\tfrac{1}{2}(t - x_i)^2 + \eta\,\mathbf{1}_{\{t\neq 0\}}\right\} \quad \text{for } i = 1, \ldots, n.$$

Consider the function

$$\phi_i(t) = \tfrac{1}{2}(t - x_i)^2 + \eta\,\mathbf{1}_{\{t\neq 0\}}.$$

There are two cases:

- If we restrict to $t \neq 0$, the quadratic term is minimized at $t = x_i$, giving value $\phi_i(x_i) = \eta$.
- At $t = 0$, the value is $\phi_i(0) = \frac{1}{2}x_i^2$.

Therefore,

$$u_i = \begin{cases} 0, & \text{if } \frac{1}{2}x_i^2 \leq \eta \ \text{ (i.e., } |x_i| \leq \sqrt{2\eta}), \\ x_i, & \text{if } \frac{1}{2}x_i^2 > \eta \ \text{ (i.e., } |x_i| > \sqrt{2\eta}). \end{cases}$$

Coordinatewise, we obtain the *hard-thresholding* operator at level $\sqrt{2\eta}$:

$$\big(\mathrm{prox}_{\eta\|\cdot\|_0}(x)\big)_i = H_{\sqrt{2\eta}}(x_i) := \begin{cases} x_i, & |x_i| > \sqrt{2\eta}, \\ 0, & |x_i| \leq \sqrt{2\eta}. \end{cases}$$

At the boundary $|x_i| = \sqrt{2\eta}$, both $t = 0$ and $t = x_i$ have equal cost η, so the prox is set-valued. It is customary to choose the zero solution (the expression above), yielding a single-valued operator.