Optimization and Algorithms 2023/24

Instituto Superior Técnico

1^{st} exam – November 9, 2023

Student ID: __________ **Name:** ____

- Exam duration 1h45
- Mark the answers to multiple-choice questions in the table below. For a question worth C points with n options to choose from, you will lose $C/(n-1)$ points if you answer it incorrectly. There is no penalty for leaving it blank
- At the end of the exam you should turn in the table of answers to multiple-choice questions, as well as the answers to open-ended problems **in separate sheets**
- Make sure that you fill out your **name and student ID in all sheets**
- **Justify** your answers to all open-ended problems. Portuguese-speaking students should answer in **Portuguese**, others in **English**

Answers to multiple-choice questions

	a	b	c	d	e	f	g	h	i
Q1									
Q2									
Q3									
Q4									

Q 1. [1.5 val] At the k-th iteration of the Levenberg-Marquardt algorithm the residual $r(x_k)$ and its Jacobian $J_r(x_k)$ are given by

$$r(x_k) = \begin{bmatrix} 0.11 \\ -0.26 \\ 0.34 \end{bmatrix}, \qquad J_r(x_k) = \begin{bmatrix} 0.22 & -0.31 \\ 0.09 & 0 \\ 0 & 0.10 \end{bmatrix}$$

If the stopping criterion is based on the norm of the gradient of the cost function $f(x) = \frac{1}{2}\|r(x)\|^2$, choose the minimum value for the tolerance ϵ such that $\|\nabla f(x_k)\| \le \epsilon$ will be met at the current iteration, causing the algorithm to stop.

a) 10^{-4} **b)** 10^{-3} **c)** 10^{-2} **d)** 10^{-1} **e)** 1 **f)** 10

Q 2. [1.5 val] Choose the smallest value of α for which the following function becomes strongly convex

$$f(x,y) = (4y - 5x + 1)_+ + \alpha x^2 + 4xy + 3y^2$$

a) 1 **b)** 2 **c)** 3 **d)** 4 **e)** 5 **f)** 6

Q 3. [1.5 val] A satellite orbits the earth in circles along the equatorial plane. In that plane, using an appropriate earth-centered inertial frame, denote the satellite's position at time t by $x(t) \in \mathbb{R}^2$ and its velocity by $v(t) \in \mathbb{R}^2$. The satellite's acceleration results from the combined actions of the thruster, which applies a force $u(t) \in \mathbb{R}^2$, and gravity, which pulls it towards the center of the earth with a force that is inversely proportional to $\|x(t)\|^2$. The (very) simplified and discretized equations of motion are

$$x(t+1) = x(t) + \gamma v(t), \qquad\qquad v(t+1) = v(t) + \alpha u(t) - \beta \frac{x(t)}{\|x(t)\|^3},$$

where α, β, γ are constants.

From an initial state x_{ini}, v_{ini} at time $t = 1$, we wish to optimally design a sequence of T inputs $u(t)$, $t = 1, \dots, T$ such that after the last one the satellite reaches a target orbit with a specified radius, $\|x(T+1)\| = H$, and speed, $\|v(T+1)\| = V$, the latter being purely tangential, $x(T+1)^T v(T+1) = 0$. Throughout the manoeuvre the satellite should remain above a minimum safe radius S. As fuel consumption is proportional to the sum of $\|u(t)\|^2$ over $t = 1, \dots, T$, the following formulation is proposed to design the most energetically efficient actuation

$$\begin{array}{ll} \underset{\substack{u(t), x(t), v(t),\\ t=1,\dots,T}}{\text{minimize}} & \frac{1}{2}\sum_{t=1}^{T}\|u(t)\|^2 + \frac{\rho}{2}\cdot \boxed{\quad ② \quad} + \frac{\rho}{2}(\|v(T+1)\| - V)^2 \\ \text{subject to} & x(1) = x_{\text{ini}},\ v(1) = v_{\text{ini}}, \\ & x(t+1) = x(t) + \gamma v(t), \\ & v(t+1) = v(t) + \alpha u(t) - \beta \frac{x(t)}{\|x(t)\|^3}, \quad t = 1, \dots, T, \\ & \boxed{\quad ① \quad}, \\ & \|x(t)\| \geq S, \quad t = 1, \dots T \end{array}$$

where $\rho > 0$ is a regularization parameter. Choose below the content for box ②.

a) $\|(I - v(T+1)v(T+1)^T)x(T+1)\| = 0$ **b)** $\sum_{t=1}^{T+1}(\|x(t)\| - H)^2$
c) $x(T+1)^T v(T+1)$ **d)** $(\|x(T+1)\| - H)^2$
e) $\|x(T+1)\| \geq H$ **f)** $x(T+1)^T v(T+1) = 0$

Q 4. [1.5 val] Consider the constrained optimization problem

$$\begin{array}{ll} \underset{x,\,y}{\text{minimize}} & (x-1)^2 + 2(y-2)^2 \\ \text{subject to} & x = 2y \end{array}$$

whose optimal solution is $(x^*, y^*) = (2, 1)$. Use the penalty method, with standard penalization and penalty parameter $c = 1$, to build a cost function from the given cost and constraint functions. As you know, the unconstrained minimum of this penalized cost function approximates (x^*, y^*). Express this function as a regular quadratic and select below the matrix A for the term $\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$.

a) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ **b)** $\begin{bmatrix} 4 & -6 \\ -6 & 14 \end{bmatrix}$ **c)** $\begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix}$ **d)** $\begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$ **e)** $\begin{bmatrix} 2 & -2 \\ -8 & 1 \end{bmatrix}$ **f)** $\begin{bmatrix} 3 & -4 \\ -4 & 10 \end{bmatrix}$

P 1. [4.5 val] We wish to determine the largest circle that will fit inside a polygon. The specific polygon that we consider here is an isosceles right triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$.

a) **[1 val]** The triangle can be described as the intersection of a set of half-spaces, each encoded as a linear inequality $\{x \in \mathbb{R}^2 : a_i^T x \leq b_i\}$. Find the set of a_i and b_i.

b) **[1.5 val]** Argue that (i) a linear function $a_i^T x$ can take on a maximum value of $r\|a_i\|$ when x may be any point inside a circle of radius r centered at the origin. Also, (ii) for any x inside a circle of radius r centered at c, the same linear function takes on a maximum of $a_i^T c + r\|a_i\|$.

c) **[0.5 val]** From the previous result, conclude that a circle of radius r centered at c lies entirely on a halfspace defined by $a_i^T x \leq b_i$ if $a_i^T c + r\|a_i\| \leq b_i$.

d) **[1.5 val]** Formulate the optimization problem for finding c and r for the largest circle contained in the given polygon. The optimal value c^* is called the *Chebyshev center* of the polygon.

P 2. [4.5 val] In range-based source localization the source position may be found by solving the convex optimization problem

$$\underset{x}{\text{minimize}} \quad \sum_{k=1}^{M} ((\|x - a_k\| - r_k)_+)^2$$

where a_k denotes the position of the k-th reference point (anchor), r_k is the corresponding range measurement, and $(\cdot)_+$ denotes $\max\{0, \cdot\}$ as usual.

a) **[1.5 val]** Argue that each term $((\|x - a_k\| - r_k)_+)^2$ may be replaced by the optimal value of the subproblem

$$\begin{array}{ll} \underset{y_k}{\text{minimize}} & \|x - (a_k + y_k)\|^2 \\ \text{subject to} & \|y_k\| \leq r_k \end{array}$$

In words, given x find the closest point to it on a circle with radius r_k centered in a_k and evaluate the squared distance between them. Discuss both cases when x is located outside or inside the circle.

b) **[1.5 val]** If x were known to be located outside the circle with radius r_k centered on a_k, then the inequality $\|y_k\| \leq r_k$ could be replaced with an equality constraint $\|y_k\|^2 = r_k^2$. Under that assumption write the KKT equations and solve them, confirming the intuitive solution for y_k that you gave to the previous question.

c) **[1.5 val]** From the previous discussion, the full relaxed source localization problem is equivalently written as

$$\begin{array}{ll} \underset{x, y_1, \ldots, y_M}{\text{minimize}} & \sum_{k=1}^{M} \|x - (a_k + y_k)\|^2 \\ \text{subject to} & \|y_k\| \leq r_k, \quad k = 1, \ldots, M \end{array}$$

This formulation defines x as the solution of a simple least-squares problem if the variables $y_1, \ldots, y_M$ are regarded as fixed. Find the optimal x as a function of $y_1, \ldots, y_M$, which would allow us to eliminate it from the problem, retaining only the y_k as variables.

(turn to the last page)

P 3. **[5 val]** The log-sum-exp function (here in 2D, for simplicity) $f(x, y) = \log(e^x + e^y)$ is often used as a differentiable surrogate for $\max(x, y)$.

a) **[1 val]** Draw the contour lines of the related function $\{(x, y) \in \mathbb{R}^2 : \|(x, y)\|_\infty \triangleq \max(|x|, |y|) = c\}$ for $c = 1, 2,$ and 3.

b) **[1.5 val]** Show that the gradient and Hessian of $f(x, y)$ are given by

$$\nabla f(x, y) = \frac{1}{e^x + e^y} \begin{bmatrix} e^x \\ e^y \end{bmatrix}, \qquad \nabla^2 f(x, y) = \frac{e^x e^y}{(e^x + e^y)^2} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix}$$

c) **[1 val]** From the previous result, what can you conclude regarding the convexity of f? Given the sketch that you made for the contour lines of $\|(x, y)\|_\infty$, which coincides with $\max(x,\, y)$ on the first quadrant, is it surprising that the Hessian of f is not full rank?

d) **[1.5 val]** Show that, for $z \in \mathbb{R}^n$, the function defined as

$$g(z) = \log\left(e^{s_1^T z + r_1} + e^{s_2^T z + r_2}\right)$$

is convex for arbitrary $s_i \in \mathbb{R}^n$, $r_i \in \mathbb{R}$, $i = 1, 2$.

Optimization and Algorithms 2023/24

Instituto Superior Técnico

Answers to the 1st exam – November 9, 2023

- Answers are only schematically presented, as they are meant to complement the students' own solutions to the problems

---

Answers to multiple-choice questions

	a	b	c	d	e	f	g	h	i
Q1		✓							
Q2		✓							
Q3				✓					
Q4			✓						

Q 1. At the k-th iteration of the Levenberg-Marquardt algorithm the residual $r(x_k)$ and its Jacobian $J_r(x_k)$ are given by

$$r(x_k) = \begin{bmatrix} 0.11 \\ -0.26 \\ 0.34 \end{bmatrix}, \qquad J_r(x_k) = \begin{bmatrix} 0.22 & -0.31 \\ 0.09 & 0 \\ 0 & 0.10 \end{bmatrix}$$

If the stopping criterion is based on the norm of the gradient of the cost function $f(x) = \frac{1}{2}\|r(x)\|^2$, choose the minimum value for the tolerance ϵ such that $\|\nabla f(x_k)\| \leq \epsilon$ will be met at the current iteration, causing the algorithm to stop.

a) 10^{-4} **b)** 10^{-3} **c)** 10^{-2} **d)** 10^{-1} **e)** 1 **f)** 10

Note: We have $\nabla f(x_k) = J_r(x_k)^T r(x_k) = \begin{bmatrix} 8 \times 10^{-4} \\ -10^{-4} \end{bmatrix}$ and $\|\nabla f x_k\| \approx 8 \times 10^{-4}$. Any ϵ greater than this will terminate the iterative Levanberg-Marquardt algorithm. The smallest such value in the list of options is $\epsilon = 10^{-3}$.

Q 2. Choose the smallest value of α for which the following function becomes strongly convex

$$f(x, y) = (4y - 5x + 1)_+ + \alpha x^2 + 4xy + 3y^2$$

a) 1 **b)** 2 **c)** 3 **d)** 4 **e)** 5 **f)** 6

Note: f is the sum of a convex function $(4y-5x+1)_+$ and a quadratic $\alpha x^2 + 4xy + 3y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \alpha & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$. This matrix will be positive definite for $3\alpha - 4 > 0$, making the quadratic form strongly convex. Adding to it the convex function $(4y - 5x + 1)_+$ will preserve the strong convexity in f. Note that $(4y-5x+1)_+$ has zero curvature everywhere, and is in fact zero on the halfspace $4y - 5x + 1 \leq 0$, so there is no way to add it to a quadratic and obtain a function f with nonvanishing curvature everywhere (i.e., strong convexity) unless the quadratic itself is strongly convex.

Q 3. A satellite orbits the earth in circles along the equatorial plane. In that plane, using an appropriate earth-centered inertial frame, denote the satellite's position at time t by $x(t) \in \mathbb{R}^2$ and its velocity by $v(t) \in \mathbb{R}^2$. The satellite's acceleration results from the combined actions of the thruster, which applies a force $u(t) \in \mathbb{R}^2$, and gravity, which pulls it towards the center of the earth with a force that is inversely proportional to $\|x(t)\|^2$. The (very) simplified and discretized equations of motion are

$$x(t+1) = x(t) + \gamma v(t), \qquad\qquad v(t+1) = v(t) + \alpha u(t) - \beta \frac{x(t)}{\|x(t)\|^3},$$

where α, β, γ are constants.

From an initial state x_{ini}, v_{ini} at time $t = 1$, we wish to optimally design a sequence of T inputs $u(t)$, $t = 1, \ldots, T$ such that after the last one the satellite reaches a target orbit with a specified radius, $\|x(T+1)\| = H$, and speed, $\|v(T+1)\| = V$, the latter being purely tangential, $x(T+1)^T v(T+1) = 0$. Throughout the manoeuvre the satellite should remain above a minimum safe radius S. As fuel consumption is proportional to the sum of $\|u(t)\|^2$ over $t = 1, \ldots, T$, the following formulation is proposed to design the most energetically efficient actuation

$$\begin{array}{ll} \underset{\substack{u(t),x(t),v(t),\\ t=1,\ldots,T}}{\text{minimize}} & \frac{1}{2}\sum_{t=1}^{T}\|u(t)\|^2 + \frac{\rho}{2}\cdot \boxed{\quad ② \quad} + \frac{\rho}{2}(\|v(T+1)\| - V)^2 \\ \text{subject to} & x(1) = x_{\text{ini}},\ v(1) = v_{\text{ini}}, \\ & x(t+1) = x(t) + \gamma v(t), \\ & v(t+1) = v(t) + \alpha u(t) - \beta \frac{x(t)}{\|x(t)\|^3}, \quad t = 1, \ldots, T, \\ & \boxed{\quad ① \quad}, \\ & \|x(t)\| \geq S, \quad t = 1, \ldots T \end{array}$$

where $\rho > 0$ is a regularization parameter. Choose below the content for box ②.

a) $\|(I - v(T+1)v(T+1)^T)x(T+1)\| = 0$ **b)** $\sum_{t=1}^{T+1}(\|x(t)\| - H)^2$

c) $x(T+1)^T v(T+1)$ **d)** $(\|x(T+1)\| - H)^2$

e) $\|x(T+1)\| \geq H$ **f)** $x(T+1)^T v(T+1) = 0$

Q 4. Consider the constrained optimization problem

$$\begin{array}{ll} \underset{x,\,y}{\text{minimize}} & (x-1)^2 + 2(y-2)^2 \\ \text{subject to} & x = 2y \end{array}$$

whose optimal solution is $(x^*, y^*) = (2, 1)$. Use the penalty method, with standard penalization and penalty parameter $c = 1$, to build a cost function from the given cost and constraint functions. As you know, the unconstrained minimum of this penalized cost function approximates (x^*, y^*). Express this function as a regular quadratic and select below the matrix A for the term $\begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix}$.

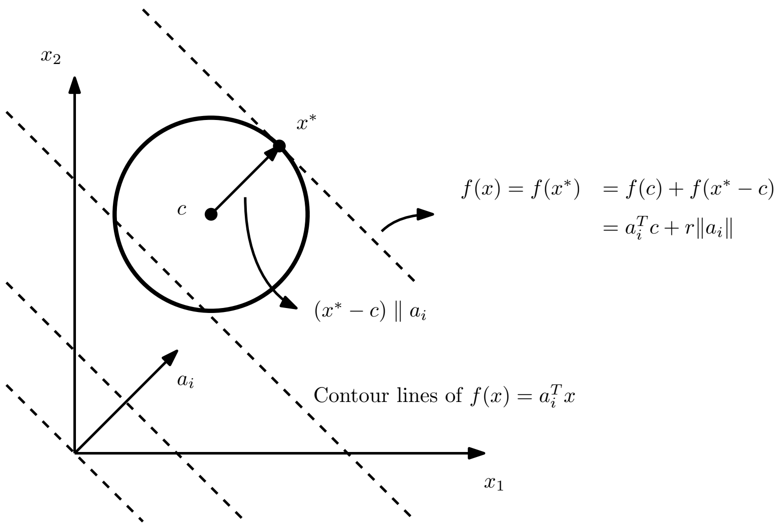
a) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ **b)** $\begin{bmatrix} 4 & -6 \\ -6 & 14 \end{bmatrix}$ **c)** $\begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix}$ **d)** $\begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$ **e)** $\begin{bmatrix} 2 & -2 \\ -8 & 1 \end{bmatrix}$ **f)** $\begin{bmatrix} 3 & -4 \\ -4 & 10 \end{bmatrix}$

Note: The penalized cost function is $(x-1)^2 + 2(y-2)^2 + c(x-2y)^2$. Expanding it, the quadratic terms (i.e., those pertaining to x^2, y^2 and xy) are $(1+c)x^2 + (2+4c)y^2 - 4cxy$, which can be grouped as $\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1+c & -2c \\ -2c & 2+4c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

P 1 (Chebyshev center of a polygon)**.** We wish to determine the largest circle that will fit inside a polygon. The specific polygon that we consider here is an isosceles right triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$.

a) The triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$ can be described as the intersection of three halfspaces defined by $x_1 \geq 0$, $x_2 \geq 0$ and $x_1 + x_2 \leq 1$. In standard form these inequalities are written as $a_i^T x \leq b_i$ with $a_1 = (-1,\ 0)$, $a_2 = (0,\ -1)$, $a_3 = (1,\ 1)$ and $b_1 = b_2 = 0$, $b_3 = 1$.

b) The contour lines of a linear function $f(x) = a_i^T x$ are parallel to each other and orthogonal to vector a_i, which points in the direction of the fastest increase of f. When x is constrained to lie on a circle of radius r centered at the origin, the maximum of f is attained when x lies as far as possible along the direction of a_i, i.e., it is located at the boundary of the circle and aligned with a_i so that $x = \frac{a_i}{\|a_i\|} r$. The value of the function is $f(x) = \frac{a_i^T a_i}{\|a_i\|} r = r\|a_i\|$. More compactly, the Cauchy-Schwarz inequality states that $a_i^T x \leq \|a_i\| \cdot \|x\|$, and equality is attained when x is parallel to a_i. This can be maximized by choosing x to lie on the boundary of the circle, $\|x\| = r$.
When the circle is centered at c we apply the same argument, shifting the origin to c, as $a_i^T(x - c) \leq \|a_i\| \cdot \|x - c\| \leq r\|a_i\|$, i.e., $a_i^T x \leq a_i^T c + r\|a_i\|$. Equality is attained when x lies at the boundary of the circle such that $x - c$ is aligned with a_i. This is illustrated in the following figure



c) By definition of maximum we have that, for any x lying on a circle of radius r and center c, $a_i^T x \leq a_i^T c + r\|a_i\|$. If the right-hand side is $\leq b_i$, then the same must be true for the left-hand side, and $a_i^T x \leq b_i$ implies that any x on the circle lies on the halfspace.

d)

$$\begin{array}{ll} \underset{c,\, r}{\text{maximize}} & r \\ \text{subject to} & a_i^T c + r\|a_i\|, \quad i = 1, 2, 3, \\ & r \geq 0 \end{array}$$

The constraint $r \geq 0$, however, is redundant as long as the intersection of halfspaces is not empty.

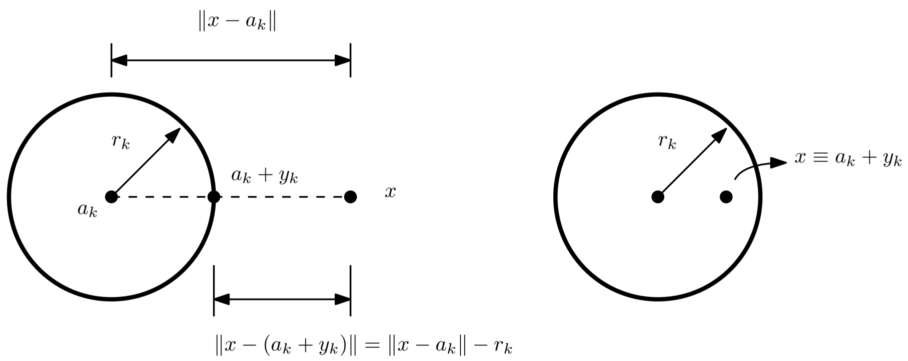
P 2 (Range-based source localization)**.** Baseline convex optimization problem for range-based source localization

$$\underset{x}{\text{minimize}} \quad \sum_{k=1}^{M} ((\|x - a_k\| - r_k)_+)^2$$

a) We replace each term $((\|x - a_k\| - r_k)_+)^2$ with the optimal value of the subproblem:

$$\begin{aligned} &\underset{y_k}{\text{minimize}} && \|x - (a_k + y_k)\|^2 \\ &\text{subject to} && \|y_k\| \le r_k \end{aligned}$$

The figure below illustrates the concept.



If x is located outside the circle with radius r_k centered in a_k, then $\|x - a_k\| > r_k$ and the closest point to x on the circle, $a_k + y_k$, will be located on the boundary. As shown in the figure, the distance between x and the optimal $a_k + y_k$ is $\|x - (a_k + y_k)\| = \|x - a_k\| - r_k > 0$, and we thus have $\|x - (a_k + y_k)\|^2 = (\|x - a_k\| - r_k)^2 = ((\|x - a_k\| - r_k)_+)^2$ as intended. When x is located inside the circle $\|x - a_k\| < r_k$ and we can choose y_k such that $y_k + a_k = x$. Then $\|x - (a_k + y_k)\| = 0 = (\|x - a_k\| - r_k)_+$.

b) For cost and constraint functions $f(y_k) = \|x - (a_k + y_k)\|^2$, $h(y_k) = \|y_k\|^2 - r_k^2$ we have the KKT conditions:

$$\begin{cases} \nabla f(y_k) + \lambda \nabla h(y_k) = 0 \\ h(y_k) = 0 \end{cases} \rightarrow \begin{cases} y_k - (x - a_k) + \lambda y_k = 0 \\ \|y_k\|^2 = r_k^2 \end{cases} \rightarrow \begin{cases} y_k = \frac{x - a_k}{1+\lambda} \\ \frac{\|x - a_k\|^2}{(1+\lambda)^2} = r_k^2 \end{cases} \rightarrow$$

$$\begin{cases} y_k = \pm r_k \frac{x - a_k}{\|x - a_k\|} \\ 1 + \lambda = \pm \frac{\|x - a_k\|}{r_k} \end{cases} \rightarrow f(y_k) = \left(1 \mp \frac{r_k}{\|x - a_k\|}\right)^2 \|x - a_k\|^2$$

Because $r_k/\|x - a_k\| \ge 0$ the lowest cost is $(1 - r_k/\|x - a_k\|)^2 \|x - a_k\|^2$, obtained for $y_k = r_k \frac{x - a_k}{\|x - a_k\|}$, which agrees with the left-side figure above.

c)

$$\underset{x}{\text{minimize}} \quad \sum_{k=1}^{M} \|x - (a_k + y_k)\|^2 = \|Ax - b\|^2$$

with

$$A = \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix}, \qquad b = \begin{bmatrix} a_1 + y_1 \\ \vdots \\ a_M + y_M \end{bmatrix}$$

Since A has full column rank the standard least-squares solution is

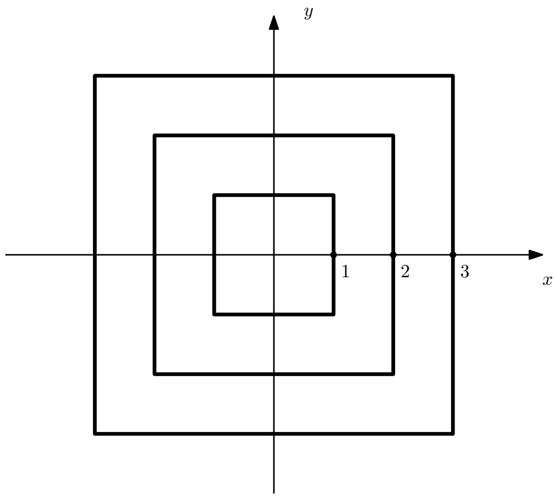
$$x = (A^T A)^{-1} A^T b = \frac{\sum_{k=1}^{M} a_k + y_k}{M}$$

which is recognized as the centroid of all the points $a_k + y_k$.

P 3 (Properties of log-sum-exp)**.**

$$f(x, y) = \log\big(e^x + e^y\big) \approx \max(x, y)$$

a) Contour lines for the l_∞ norm are squares whose sides are parallel to the coordinate axes. This function has linear slope everywhere.



b)

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \frac{1}{e^x + e^y} \begin{bmatrix} e^x \\ e^y \end{bmatrix},$$

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \frac{e^x e^y}{(e^x + e^y)^2} \underbrace{\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}}_{\begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix}}$$

c) For any vector $v = (v_1, v_2)$ we have $v^T \nabla^2 f(x,y) v = \frac{e^x e^y}{(e^x + e^y)^2}(v_1 - v_2)^2 \geq 0$, which proves that the Hessian is always positive semidefinite. It follows that f is convex. Because $f(x,y)$ approximates $\max(x,y)$ and the latter is a function with linear slope, one would not expect $\nabla^2 f(x,y)$ to be positive *definite* (hence full rank and invertible), otherwise f would have nonvanishing curvature everywhere, unlike max. The non-invertibility of $\nabla^2 f(x,y)$ is obvious from the given rank-1 factorization ($\nabla^2 f(x,y) v = 0$ when $v_1 = v_2$) and from its eigenvalues (which include $\lambda = 0$).

d) $g(z) = \log\big(e^{s_1^T z + r_1} + e^{s_2^T z + r_2}\big)$, with $z \in \mathbb{R}^n$, is the composition of the convex 2D log-sum-exp function $f(x,y)$ with an affine map from $\mathbb{R}^n$ to $\mathbb{R}^2$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} s_1^T \\ s_2^T \end{bmatrix}}_{S \in \mathbb{R}^{2 \times n}} z + \underbrace{\begin{bmatrix} r_1 \\ r_2 \end{bmatrix}}_{r \in \mathbb{R}^2}$$

and is therefore convex.