CE 5999: Convex Optimization **Midterm Exam** **Spring 2023**

Your Name:

Your Signature:

- **Exam duration:** 2 hours and 1 minute.
- This exam is closed book, closed notes, closed laptops, closed phones, closed tablets, closed pretty much everything, besides a one-page of formula sheet.
- Solve whatever problems you can solve. The exam is long but (mostly) straightforward. If you've been studying, you'll find it a breather. If you haven't, then I hope the agony is bearable.
- **No calculators** of any kind are allowed.
- In order to receive credit, you must **show all of your work**.
- If you need more room, use the back of the pages and indicate that you have done so.
- This exam has 12 pages, plus this cover sheet. Please make sure that your exam is complete, that you read all the exam directions and rules. Good luck! I hope you do well.

Question Number	Maximum Points	Your Score
1	20	
2	65	
3	25	
4	20	
5	30	
6	40	
Total	200	

1. (20 total points) In the following optimization problem, $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$ are given, and $x \in \mathbb{R}^n$ is the optimization variable.

$$\begin{array}{ll} \text{minimize} & e^x + 2e^{-2x} \\ \text{subject to} & \|x - a\|_2 \geq \|x - b\|_2 \end{array}$$

   (a) (20 points) Determine if this problem can be written as a convex optimization problem.

Solution: The function e^{-2x} is convex as composition of the exponential with a linear function. The objective is convex as nonnegative sum of convex functions.

The constraint can be equivalently squared, which gives

$$\begin{aligned} \|x - a\|_2^2 \geq \|x - b\|_2^2 &\Leftrightarrow x^T x - 2a^T x + a^T a \geq x^T x - 2b^T x + b^T b \\ &\Leftrightarrow 2(b - a)^T x \geq -a^T a + b^T b \end{aligned}$$

which defines a halfspace.

Therefore, the problem is convex.

2. (65 total points) Answer the following miscellaneous questions.

   (a) (10 points) Represent the following inequality as a linear matrix inequality, given that A, b, θ are given quantities:

$$||Ax - b|| \leq \theta.$$

Solution. The above inequality is equivalent to $||Ax - b||^2 \leq \theta^2$ which can be written as

$$\theta^2 - (Ax - b)^T (Ax - b) \geq 0.$$

Using Schur complement, we can write this as

$$\begin{bmatrix} I & Ax - b \\ (Ax - b)^T & \theta^2 \end{bmatrix} \succeq 0$$

Et voila! This is an LMI.

(b) (10 points) Let $\mathcal{S}_+^n$ be the set of positive semi-definite matrices of size n. Prove that $\mathcal{S}$ is a convex set.

Solution. Define two matrices in $\mathcal{S}_+^n$, i.e., $X, Y \in \mathcal{S}_+^n$. You simple need to prove that the convex combination of X and Y, $\alpha X + (1-\alpha)Y$ is also in $\mathcal{S}_+^n$. Well, since $\alpha > 0$ (or between one and zero), then the matrix $Z = \alpha X + (1-\alpha)Y$ will also be a positive definite matrix with non-negative eigenvalues, hence the set $\mathcal{S}_+^n$ is indeed convex.

(c) (10 points) Is the function

$$f(x_1, x_2) = \frac{1}{(x_1-2)^2 + (x_2+1)^2 + 3}$$

locally convex, concave, or neither in the neighborhood of the point $x^{(0)} = [2 \ -1]^T$?

Solution. For this problem, we can compute the Hessian matrix and evaluate it at the operating point, which returns

$$\nabla^2 f(x^{(0)}) = \begin{bmatrix} -2/9 & 0 \\ 0 & -2/9 \end{bmatrix}.$$

This means that the function around the prescribed operating point is concave. Notice that although the function form and its gradient and Hessian are seemingly cumbersome to compute, most of the terms are equal to zero since we're evaluating the function at a specific operating point that nullifies a lot of terms.

(d) (10 points) Is the quadratic form

$$f = f(x_1, x_2, x_3) = x^T \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} x$$

positive definite, negative semidefinite, or indefinite? Justify your answer carefully.

Solution. First, we'll have to transform the above quadratic form into a symmetric quadratic one. To do, we can write

$$f = f(x_1, x_2, x_3) = x^T \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} x = \frac{1}{2} x^T \left(\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}^T \right) x$$

which is equal to

$$f = x^T \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} x$$

This matrix has positive and negative and zero principal minors, meaning that this quadratic form is actually indefinite. Also, you can verify that by checking the eigenvalues of A (one is zero, one is negative, and another is positive).

(e) (15 points) Consider the problem

$$\begin{array}{ll} \text{minimize} & 2x_1^2 + 3x_2^2 - 4x_1 + 6x_2 + 11 \\ \text{subject to} & x_1 \geq 0,\, x_2 \geq 0. \end{array}$$

Determine if the problem is convex and show that its solution is the point (1,0).

Solution: The gradient vector is

$$\nabla f(x) = \begin{bmatrix} 4x_1 - 4 \\ 6x_2 + 6 \end{bmatrix}.$$

The problem is convex, therefore, the necessary and sufficient optimality condition is $\nabla f(x^*)^T(x - x^*)$ for all $x \succeq 0$. Substituting $x_1^* = 1$ and $x_2^* = 0$ we obtain

$$\begin{bmatrix} 0 \\ 6 \end{bmatrix}^T \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} = 6x_2 \geq 0$$

which holds for all $x_2 \geq 0$.

(f) (10 points) Determine if the function $f(x_1, x_2) = 2x_1^3 - 3x_2^2$ is convex or concave. Then investigate whether there is a subset of $\mathbb{R}^2$ over which the function is convex or concave.

Solution: The gradient vector and Hessian matrix are

$$\nabla f(x) = \begin{bmatrix} 6x_1^2 \\ -6x_2 \end{bmatrix}, \quad \nabla^2 f(x) = \begin{bmatrix} 12x_1 & 0 \\ 0 & -6 \end{bmatrix}.$$

Clearly, the function is neither concave nor convex. However, for $x_1 < 0$ and $x_2 \in \mathbb{R}$, the function is concave as the Hessian matrix in negative definite.

3. (25 total points) The objective of this problem is to show you how LMIs are nothing but nonlinear (but still convex) optimization problems. You are given the following optimization problem:

$$\begin{aligned} \textbf{OP1:}\quad & \text{minimize} && \textbf{trace}(P) = p_1 + p_3 x` \\ & \text{subject to} && AP + PA^\top + Q = 0 && (1) \\ & && P = P^\top \succ 0 && (2) \end{aligned}$$

where $A = \begin{bmatrix} -1 & -2 \\ 0 & -2 \end{bmatrix}$ and $Q = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. For the above problem, assume that $P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$ is the optimization variable. In other words, you have three variables to solve for, since P is symmetric and positive definite.

(a) (10 points) Define a new variable $x = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}^\top$ and write the first constraint in **OP1** as a linear system of equations, i.e., $\tilde{A}x = b$, where $\tilde{A} \in \mathbb{R}^{4\times 3}$ and $b \in \mathbb{R}^{4\times 1}$ are matrices you should determine.

Solution. We proceed by writing down the first constraint for the given matrices A and Q. This yields

$$\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} -2x_1 - 4x_2 & -3x_2 - 2x_3 \\ -3x_2 - 2x_3 & -4x_3 \end{bmatrix}$$

which yields a system of linear equations with three unknowns as:

$$\tilde{A}x = b,\ \tilde{A} = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix},\ b = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}.$$

(b) (5 points) Write the second positive definiteness constraint on P (i.e., $P = P^\top \succ 0$) as a nonlinear set of equations. Remember that a matrix is positive definite if and only if all of its leading principal minors are positive. You should obtain two inequality constraints here.

Solution. This constraint can simple be written as $x_1 > 0, x_1 x_3 - x_2 > 0$

(c) (10 points) Using the above transformations, write **OP1** as a simple optimization problem with a linear cost function, linear equality constraints, and quadratic inequality constraints. You should get something like this:

$$\begin{aligned} \textbf{OP2} \equiv \textbf{OP1} \quad \underset{x}{\text{minimize}} \quad & c^\top x \\ \text{subject to} \quad & \tilde{A}x = b && (3) \\ & x_1 > 0 && (4) \\ & x^\top Q x + x^\top \tilde{b} + c > 0 && (5) \end{aligned}$$

where $c, \tilde{A}, b, Q, \tilde{b}$, and c are constant matrices and vectors that you should have already determined.

Solution. In this formulation, since we're minimizing the trace, then $c = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^\top$. One can also see that

$$\tilde{b} = 0, c = 0, \tilde{Q} = \begin{bmatrix} 0 & 0 & 0.5 \\ 0 & -1 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}.$$

4. (20 points) Consider the following problem where $A_i \in \mathbb{R}^{m_i \times n}$, $b_i \in \mathbb{R}^{m_i}$, $\lambda_i > 0$ $(i = 1, 2)$, are given, and $x \in \mathbb{R}^n$ is the optimization variable.

$$\text{minimize} \;\; \lambda_1 \|A_1 x - b_1\|_2^2 + \lambda_2 \|A_2 x - b_2\|_2^2$$

   (a) (20 points) Determine if the problem is convex and find a closed-form solution. You may assume that matrix $\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2$ is invertible.

Solution. The objective function is written as follows:

$$\begin{aligned} f(x) &= \lambda_1 \|A_1 x - b_1\|_2^2 + \lambda_2 \|A_2 x - b_2\|_2^2 \\ &= \lambda_1 (A_1 x - b_1)^T (A_1 x - b_1) + \lambda_2 (A_2 x - b_2)^T (A_2 x - b_2) \\ &= \lambda_1 \left(x^T A_1^T A_1 x - 2 b_1^T A_1 x + b_1^T b_1 \right) + \lambda_2 \left(x^T A_2^T A_2 x - 2 b_2^T A_2 x + b_2^T b_2 \right) \\ &= x^T (\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2) x - 2(\lambda_1 b_1^T A_1 + \lambda_2 b_2^T A_2) x + \lambda_1 b_1^T b_1 + \lambda_2 b_2^T b_2. \end{aligned}$$

The objective function is a quadratic. To see that it is convex, notice first that $A_1^T A_1$ and $A_2^T A_2$ are always positive semidefinite. Their weighted sum is also positive semidefinite, using the definition. That is, for any $x \in \mathbb{R}^n$, it holds that $x^T (\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2) x \geq 0$, because $\lambda_1 x^T (A_1^T A_1) x \geq 0$ and $\lambda_2 x^T (A_2^T A_2) x \geq 0$ for all non-negative $\lambda_{1,2}$.

Therefore, the problem is convex. Setting the gradient equal to zero is a sufficient condition for optimality.

$$\nabla f(x) = 0 \Rightarrow 2(\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2) x - 2(\lambda_1 A_1^T b_1 + \lambda_2 A_2^T b_2) = 0$$

Is has been assumed that $\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2$ is full-rank, i.e., it has no zero eigenvalue. Given that it is positive semidefinite, it follows that it is in fact positive definite (and thus invertible). Solving the previous equation for x yields the optimal solution:

$$x^* = (\lambda_1 A_1^T A_1 + \lambda_2 A_2^T A_2)^{-1} (\lambda_1 A_1^T b_1 + \lambda_2 A_2^T b_2).$$

5. (30 points) Consider the following measurement model

$$y(t) = a^T(t)x + w(t)$$

where $y(t) \in \mathbb{R}$ defines the measurement data, $w(t) \in \mathbb{R}$ depicts measurement noise, vector $a \in \mathbb{R}^n$ is a constant, pre-determined vector, and $x \in \mathbb{R}^n$ is a vector parameter to be estimated. The measurements are defined for $t = 1, 2, \ldots, m$.

(a) (15 points) Assume that the probability density function of $w(t)$ is given as a Gaussian PDF of

$$p(w) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{w-\mu}{\sigma}\right)^2}$$

where $\mu = 0$ defines the mean and σ the standard deviation. Consider now that our objective is to find a maximum likelihood (ML) estimate for vector x given the m measurements. This problem can be posed as

$$\text{maximize}_{x_{\text{ML}} \in \mathbb{R}^n} \ \sum_{t=1}^{m} \log p(y(t); x) = \sum_{t=1}^{m} \log p(y(t) - a^\top(t)x)$$

Derive the solution to the optimization problem given the above assumptions and show that the ML estimate under Gaussian noise is indeed the solution to a least squares minimization problem.

Solution. Assuming that the mean is equal to zero, the optimization formulation can be written as

$$\text{maximize}_x \sum_{t=1}^{m} \log p(y(t) - a^\top(t)x) = \sum_{t=1}^{m} \log \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y(t) - a^\top(t)x}{\sigma}\right)^2} \right]$$

which is equivalent to

$$\max_x \ -\frac{m}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}||Ax - y||_2^2 \quad \equiv \quad \min_x ||Ax - y||_2^2$$

where A is the matrix with rows $a^T(1), a^T(2), \ldots, a^T(m)$. The solution to this problem is the least squares estimate

$$x_{\text{ML}} = \arg\min ||Ax - y||_2^2 = (A^T A)^{-1} A^T y.$$

(b) (15 points) What happens when the noise distribution for $w(t)$ changes to either the Laplacian distribution given by $p(w) = \frac{1}{2a} e^{-|w|/a}$ for a fixed $a > 0$ or the Uniform distribution given by $p(w) = \frac{1}{2a}$ on $[-a, a]$? Derive the maximum likelihood estimates for these two distributions too.

Solution. Using a similar derivation, we obtain

- $x_{\text{ML}} = \arg\min ||Ax - y||_1$ for the Laplacian distribution
- $x_{\text{ML}} = \{x \in \mathbb{R}^n$, such that $||Ax - y||_\infty < a\}$ for the Laplacian distribution

6. (40 points) Consider the following optimization problem

$$\begin{aligned} \min \quad & x^T Q x + c^T x && (6) \\ \text{subject to} \quad & a_i^T x \leq b_i, \quad i = 1, \ldots, m. && (7) \end{aligned}$$

   (a) (25 points) Assume that the scalars b_i, vector c, and matrix Q are all known. Considering now that vectors $a_i \in \mathbb{R}^n$ are all uncertain but yet remain in these ellipsoids centered around a constant $\bar{a}_i$ with orientation/size given by matrices P_i, that is:

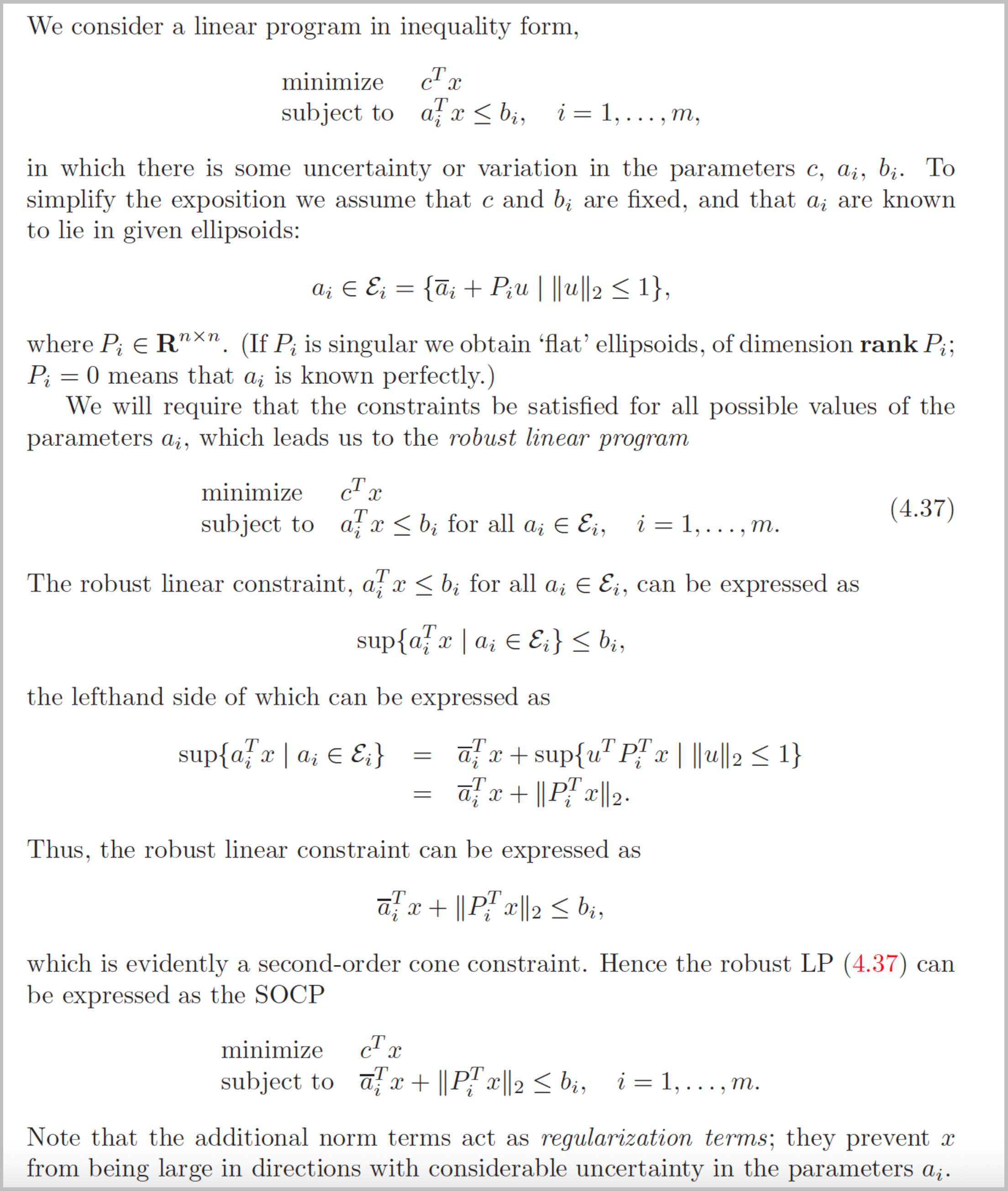
$$a_i \in \mathcal{E}_i = \{\bar{a}_i + P_i u \mid ||u||_2 \leq 1\}$$

   where matrices $P_i \in \mathbb{R}^{ntimesn}$. Notice that by setting matrix $P_i = 0$, you get a deterministic constraint, i.e., the ith constraint is certain. This form of uncertainty is called ellipsoidal uncertainty. The linear inequality constraint $a_i^T ix \leq b_i$ for all $a_i \in \mathcal{E}_i$, in this problem can be expressed as

$$\sup\{a_i^T x \mid a_i \in \mathcal{E}_i\} \leq b_i, \quad i = 1, \ldots, m.$$

   Given this, show that the above optimization problem can be written as a second order cone program (SOCP). This shows that to *robustify* LPs with some uncertain data, one obtains an SOCP.

Solution. From the textbook (Section 4.4)



We consider a linear program in inequality form,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i, \quad i = 1, \ldots, m, \end{array}$$

in which there is some uncertainty or variation in the parameters c, a_i, b_i. To simplify the exposition we assume that c and b_i are fixed, and that a_i are known to lie in given ellipsoids:

$$a_i \in \mathcal{E}_i = \{\overline{a}_i + P_i u \mid \|u\|_2 \le 1\},$$

where $P_i \in \mathbf{R}^{n \times n}$. (If P_i is singular we obtain 'flat' ellipsoids, of dimension $\textbf{rank}\, P_i$; $P_i = 0$ means that a_i is known perfectly.)

We will require that the constraints be satisfied for all possible values of the parameters a_i, which leads us to the *robust linear program*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \le b_i \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m. \end{array} \tag{4.37}$$

The robust linear constraint, $a_i^T x \le b_i$ for all $a_i \in \mathcal{E}_i$, can be expressed as

$$\sup\{a_i^T x \mid a_i \in \mathcal{E}_i\} \le b_i,$$

the lefthand side of which can be expressed as

$$\begin{array}{rcl} \sup\{a_i^T x \mid a_i \in \mathcal{E}_i\} & = & \overline{a}_i^T x + \sup\{u^T P_i^T x \mid \|u\|_2 \le 1\} \\ & = & \overline{a}_i^T x + \|P_i^T x\|_2. \end{array}$$

Thus, the robust linear constraint can be expressed as

$$\overline{a}_i^T x + \|P_i^T x\|_2 \le b_i,$$

which is evidently a second-order cone constraint. Hence the robust LP (4.37) can be expressed as the SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \overline{a}_i^T x + \|P_i^T x\|_2 \le b_i, \quad i = 1, \ldots, m. \end{array}$$

Note that the additional norm terms act as *regularization terms*; they prevent x from being large in directions with considerable uncertainty in the parameters a_i.

(b) (15 points) Rewrite the derived SOCP in the previous problem into *(i)* a semidefinite program (SDP) and *(ii)* a quadratically constrained quadratic program (QCQP), separately. What are the conditions that the problem data has to satisfy so that the QCQP formulation is convex?

Solution. The SOCP can be written as a QCQP by squaring both sides of the inequality constraint. To write the SOCP as an SDP, we note that

$$||x|| \leq t \Leftrightarrow \begin{bmatrix} tI & x \\ x^T & t \end{bmatrix} \succcurlyeq 0$$

or more generally

$$\|A_i x + b_i\|_2 \leq c_i^T x + d_i \Leftrightarrow \begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succcurlyeq 0,$$

For our problem, in the above inequality $b_i = 0$, $d_i = b_i$, $A_i = P_i^\top$, and $c_i = -\bar{a}_i^\top$.