IEOR 240, Linear Algebra Recap

Fall 2021

Overview

Vectors

We use the notation $\mathbb{R}$ for the real numbers, and $\mathbb{R}^n$ for the n-dimensional real vectors:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The space $\mathbb{R}^n$ is one where we can perform addition and scalar multiplication a way that conforms to our intuition from $\mathbb{R}^3$.

Vectors

We use the notation $\mathbb{R}$ for the real numbers, and $\mathbb{R}^n$ for the n-dimensional real vectors:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The space $\mathbb{R}^n$ is one where we can perform addition and scalar multiplication a way that conforms to our intuition from $\mathbb{R}^3$.

$$\mathbf{x}^\top = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}.$$

Matrices

We use the notation $\mathbb{R}^{m \times n}$ for the $m \times n$-real matrices,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}.$$

The space $\mathbb{R}^{m \times n}$ is one where we can perform addition and scalar multiplication a way that conforms to our intuition from $\mathbb{R}^{2 \times 2}$.

Matrices

We use the notation $\mathbb{R}^{m\times n}$ for the $m \times n$-real matrices,

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}.$$

The space $\mathbb{R}^{m\times n}$ is one where we can perform addition and scalar multiplication a way that conforms to our intuition from $\mathbb{R}^{2\times 2}$.

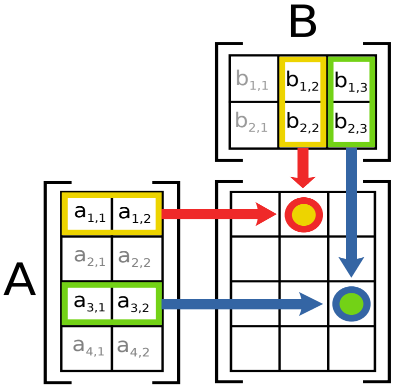
$$A^\top = \begin{bmatrix} A_{11} & \cdots & A_{m1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{mn} \end{bmatrix}.$$

Matrix Multiplication

For $A \in \mathbb{R}^{n \times k}$ and $B \in \mathbb{R}^{k \times m}$, the result of **matrix multiplication** is a matrix

$$AB = C \in \mathbb{R}^{n \times m}$$

such that $C_{ij} = \sum_{l=1}^{k} A_{il} B_{lj}$



Inner Products

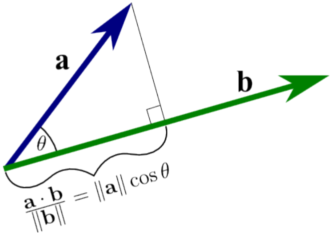
The canonical inner product on $\mathbb{R}^n$, for $a, b \in \mathbb{R}^n$:

$$\mathbf{a}^\top \mathbf{b} = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}.$$

Inner Products

The canonical inner product on $\mathbb{R}^n$, for $a, b \in \mathbb{R}^n$:

$$\mathbf{a}^\top \mathbf{b} = \sum_{i=1}^{n} a_i b_i \in \mathbb{R}.$$



Geometrically:

$$\mathbf{a}^\top \mathbf{b} = \mathbf{a} \cdot \mathbf{b} = |a||b| \cos\theta$$

Geometrical sense

Consider $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^m$. The matrix $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$

Geometrical sense

Consider $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^m$. The matrix $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$

- Columns of the matrix in multiplication on the left:

$$\mathbf{y}^\top A = \mathbf{y}^\top [\mathbf{a}_1 \ldots \mathbf{a}_m] = [\mathbf{y}^\top \mathbf{a}_1 \ \ \ldots \ \ \mathbf{y}^\top \mathbf{a}_m]$$

Geometrical sense

Consider $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^m$. The matrix $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$

- Columns of the matrix in multiplication on the left:

$$\mathbf{y}^\top A = \mathbf{y}^\top [\mathbf{a}_1 \ldots \mathbf{a}_m] = [\mathbf{y}^\top \mathbf{a}_1 \ \ \ldots \ \ \mathbf{y}^\top \mathbf{a}_m]$$

- Columns of the matrix in multiplication on the right:

$$A\mathbf{x} = [\mathbf{a}_1 \ldots \mathbf{a}_m] \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 \mathbf{a}_1 + \ldots + x_m \mathbf{a}_m = \sum_{j=1}^{m} x_j \mathbf{a}_j$$

Examples

- Vectors:

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{x}_1^\top = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad \mathbf{x}_2^\top = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}.$$

- Matrices:

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_1^\top = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 1 \end{bmatrix}.$$

Examples: inner products

- 2×3 matrix and 3-dimensional vector.

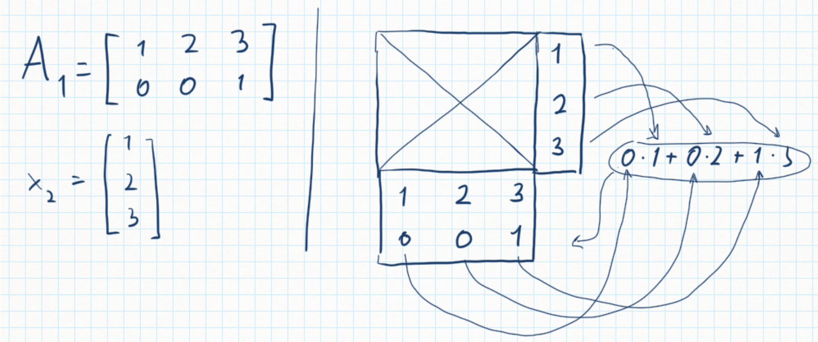
$$A_1 \mathbf{x}_2 = \begin{bmatrix} 1 \times 1 + 2 \times 2 + 3 \times 3 \\ 0 \times 1 + 0 \times 2 + 1 \times 3 \end{bmatrix} = \begin{bmatrix} 14 \\ 3 \end{bmatrix}$$

- 3×2 matrix and 2-dimensional vector.

$$A_1^\top \mathbf{x}_1 = \begin{bmatrix} 1 \times 1 + 0 \times 2 \\ 2 \times 1 + 0 \times 2 \\ 3 \times 1 + 1 \times 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}.$$

- Some other products, e.g., $x_1^\top A_1$, $x_2^\top A_1^\top$.

Examples: inner products



Inequalities notation

Consider $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$.
The i-th component of a_j is a_{ij}.

Inequalities notation

Consider $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$.
The i-th component of a_j is a_{ij}.

$$\mathbf{x} = \mathbf{1} \qquad \text{means} \qquad \begin{cases} x_1 = 1 \\ \vdots \\ x_m = 1 \end{cases}$$

Inequalities notation

Consider $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$.
The i-th component of a_j is a_{ij}.

$$\mathbf{x} = \mathbf{1} \qquad \text{means} \qquad \begin{cases} x_1 = 1 \\ \vdots \\ x_m = 1 \end{cases}$$

$$\mathbf{x} < 0 \qquad \text{means} \qquad \begin{cases} x_1 < 0 \\ \vdots \\ x_m < 0 \end{cases}$$

Inequalities notation

Consider $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ and $[\mathbf{a}_1 \ldots \mathbf{a}_m] = A \in \mathbb{R}^{n \times m}$.
The i-th component of a_j is a_{ij}.

$$\mathbf{x} = \mathbf{1} \qquad \text{means} \qquad \begin{cases} x_1 = 1 \\ \vdots \\ x_m = 1 \end{cases}$$

$$\mathbf{x} < 0 \qquad \text{means} \qquad \begin{cases} x_1 < 0 \\ \vdots \\ x_m < 0 \end{cases}$$

$$A\mathbf{x} \geq 0 \qquad \text{means} \qquad \begin{cases} \sum_{j=1}^m x_j a_{1j} \geq 0 \\ \vdots \\ \sum_{j=1}^m x_j a_{nj} \geq 0 \end{cases}$$

Arbitrage idea

The idea can be formulated as:

$$\max\{\min\{Profit\_S_1, \dots, Profit\_S_6\}\}$$

Can be re-written as:

$$\begin{aligned} \max \quad & z \\ \text{s.t.} \quad & z \leq Profit\_S_i \qquad \forall i \in \{1, \dots, 6\} \end{aligned} \tag{1}$$

General linearization of MaxMin (or MinMax)

Given that $g(y, \mathbf{x})$ is increasing with y for any $\mathbf{x} \in \mathcal{X}$,

$$\max_{\mathbf{x} \in \mathcal{X}} \quad g(\min\{f_1(\mathbf{x}), \dots, f_m(\mathbf{x})\}, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{X}} \quad & g(z, \mathbf{x}) \\ \text{s.t.} \quad & z \le f_i(\mathbf{x}) \qquad \forall i \in \{1, \dots, m\} \end{aligned}$$

General linearization of MaxMin (or MinMax)

Given that $g(y, \mathbf{x})$ is increasing with y for any $\mathbf{x} \in \mathcal{X}$,

$$\max_{\mathbf{x} \in \mathcal{X}} \quad g(\min\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\}, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{X}} \quad & g(z, \mathbf{x}) \\ \text{s.t.} \quad & z \leq f_i(\mathbf{x}) \qquad \forall i \in \{1, \ldots, m\} \end{aligned}$$

Proof by contradiction

General linearization of MaxMin (or MinMax)

Similarly, for $g(y, \mathbf{x})$ increasing with y,

$$\min_{\mathbf{x} \in \mathcal{X}} \quad g(\max\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\}, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & g(z, \mathbf{x}) \\ \text{s.t.} \quad & z \geq f_i(\mathbf{x}) \qquad \forall i \in \{1, \ldots, m\} \end{aligned}$$

General linearization of MaxMin (or MinMax)

Similarly, for $g(y, \mathbf{x})$ increasing with y,

$$\min_{\mathbf{x} \in \mathcal{X}} \quad g(\max\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\}, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & g(z, \mathbf{x}) \\ \text{s.t.} \quad & z \geq f_i(\mathbf{x}) \qquad \forall i \in \{1, \ldots, m\} \end{aligned}$$

Note also that

$$\max f(x) = -\min(-f(x))$$

Example

Suppose we have the problem:

$$\begin{aligned} \min_{x_1, x_2} \; & |x_1 + 5x_2| \\ \text{s.t.} \; & x_1 - 3x_2 \geq 2 \\ & x_1 \geq 0 \end{aligned}$$

How can we convert it into a Linear Program?

Example

Suppose we have the problem:

$$\min_{x_1,x_2} |x_1 + 5x_2|$$
$$\text{s.t. } x_1 - 3x_2 \geq 2$$
$$x_1 \geq 0$$

How can we convert it into a Linear Program?
Remember $|x| = \max\{x, -x\}$.

$$\min_{x_1,x_2} \quad \max\{x_1 + 5x_2, -x_1 - 5x_2\}$$
$$\text{s.t. } x_1 - 3x_2 \geq 2$$
$$x_1 \geq 0$$

Form 1

Create a new variable z and make:

$$\begin{aligned} \min_{x_1, x_2, z} \quad & z \\ \text{s.t.} \quad z &\geq x_1 + 5x_2 \\ z &\geq -(x_1 + 5x_2) \\ x_1 - 3x_2 &\geq 2 \\ x_1 &\geq 0 \end{aligned}$$

Form 2

Each number can be divided in its positive and negative parts:

- $35 = (35) - (0)$
- $-15 = (0) - (15)$

Or in general $x = z^+ - z^-$ where $z^+ \geq 0$ and $z^- \geq 0$.

Form 2

Using this intuition:

$$\begin{aligned} \min_{x_1, x_2, z^+, z^-} \quad & z^+ + z^- \\ \text{s.t.} \quad z^+ - z^- &= x_1 + 5x_2 \\ x_1 - 3x_2 &\geq 2 \\ x_1 \geq 0, z^+ \geq 0, z^- &\geq 0 \end{aligned}$$

Why does this work?

General linearization of Abs

Given a function $g(y, \mathbf{x})$ increasing with y for any $\mathbf{x} \in \mathcal{X}$,

$$\min_{\mathbf{x} \in \mathcal{X}} \quad g(|f(\mathbf{x})|, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & g(z^+ + z^-, \mathbf{x}) \\ \text{s.t.} \quad & z^+ - z^- = f(\mathbf{x}) \\ & z^+ \geq 0; z^- \geq 0 \end{aligned}$$

General linearization of Abs

Given a function $g(y, \mathbf{x})$ increasing with y for any $\mathbf{x} \in \mathcal{X}$,

$$\min_{\mathbf{x} \in \mathcal{X}} \quad g(|f(\mathbf{x})|, \mathbf{x})$$

is equivalent (in some sense) to

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}} \quad & g(z^{+} + z^{-}, \mathbf{x}) \\ \text{s.t.} \quad & z^{+} - z^{-} = f(\mathbf{x}) \\ & z^{+} \geq 0; z^{-} \geq 0 \end{aligned}$$

Proof by contradiction

Sketch of proof

We will show that for the optimal solution, either $z^+ = 0$ or $z^- = 0$.
Suppose $\bar{z}^+ > 0$, $\bar{z}^- > 0$ and $\bar{\mathbf{x}}$ form the optimal solution for the later problem. Then $\bar{z}^+ - \bar{z}^- = f(\bar{\mathbf{x}})$.
Introduce $a = \min\{\bar{z}^+, \bar{z}^-\}$ and $\hat{z}^+ = \bar{z}^+ - a$, $\hat{z}^- = \bar{z}^- - a$.
Observe that $(\hat{z}^+, \hat{z}^-, \bar{\mathbf{x}})$ is a feasible solution:

$$\hat{z}^+ - \hat{z}^- = \bar{z}^+ - a - \bar{z}^- + a = \bar{z}^+ - \bar{z}^- = f(\bar{\mathbf{x}})$$

and that it is better than the optimal one:

$$\bar{z}^+ + \bar{z}^- > \hat{z}^+ - \hat{z}^- \implies g(\bar{z}^+ + \bar{z}^-, \bar{\mathbf{x}}) > g(\hat{z}^+ + \hat{z}^-, \bar{\mathbf{x}}),$$

which is a contradiction with the assumption of optimality.

General Form

- **Problem:** We have the following nonlinear problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & \sum_{i=1}^{n} |x_i| \\ \text{s.t.} \quad & A\mathbf{x} = b. \end{aligned}$$

  where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

- **Description:** The objective is a non-smooth non-linear function which seems very hard to be optimized. However, this is not true!

Tricks

Two ways of linearizing the absolute value in an optimization problem:

- **Trick 1:** Introduce new variables z_i for $i = 1, 2, \ldots, n$. Add constraints $z_i \geq x_i$ and $z_i \geq -x_i$ for $i = 1, 2, \ldots, n$. Then minimize the function $\sum_{i=1}^{n} z_i$.
- **Trick 2:** Introduce new variables x_i^+ and x_i^- for $i = 1, 2, \ldots, n$. Add constraints $x_i = x_i^+ - x_i^-$ and $x_i^+ \geq 0$, $x_i^- \geq 0$ for $i = 1, 2, \ldots, n$. Then minimize the function $\sum_{i=1}^{n} \left(x_i^+ + x_i^-\right)$.

Thank you for your attention !