CS/ECE/ISYE524: Introduction to Optimization – Linear Optimization Models

Jeff Linderoth

Department of Industrial and Systems Engineering
University of Wisconsin-Madison

January 29, 2024

Matrix basics

A matrix is an array of numbers. $A \in \mathbb{R}^{m\times n}$ means that:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \qquad (m \text{ rows and } n \text{ columns})$$

Two matrices can be multiplied if inner dimensions agree:

$$\underset{(m\times p)}{C} = \underset{(m\times n)}{A}\underset{(n\times p)}{B} \qquad \text{where} \qquad c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 8 & 9 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 8 & 1\cdot 3 + 2\cdot 9 \\ 3\cdot 4 + 4\cdot 8 & 3\cdot 3 + 4\cdot 9 \\ 5\cdot 4 + 6\cdot 8 & 5\cdot 3 + 6\cdot 9 \end{bmatrix} = \begin{bmatrix} 20 & 21 \\ 44 & 45 \\ 68 & 69 \end{bmatrix}$$

Matrix basics

A matrix is an array of numbers. $A \in \mathbb{R}^{m \times n}$ means that:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \qquad (m \text{ rows and } n \text{ columns})$$

Two matrices can be multiplied if inner dimensions agree:

$$\underset{(m\times p)}{C} = \underset{(m\times n)}{A}\,\underset{(n\times p)}{B} \quad \text{where} \quad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 8 & 9 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 8 & 1\cdot 3 + 2\cdot 9 \\ 3\cdot 4 + 4\cdot 8 & 3\cdot 3 + 4\cdot 9 \\ 5\cdot 4 + 6\cdot 8 & 5\cdot 3 + 6\cdot 9 \end{bmatrix} = \begin{bmatrix} 20 & 21 \\ 44 & 45 \\ 68 & 69 \end{bmatrix}$$

Matrix basics

Transpose: The transpose operator A^T swaps rows and columns. If $A \in \mathbb{R}^{m\times n}$ then $A^\mathsf{T} \in \mathbb{R}^{n\times m}$ and $(A^\mathsf{T})_{ij} = A_{ji}$.

- $(A^\mathsf{T})^\mathsf{T} = A$
- $(AB)^\mathsf{T} = B^\mathsf{T} A^\mathsf{T}$

A vector is a column matrix. We write $x \in \mathbb{R}^n$ to mean that:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{(a vector } x \in \mathbb{R}^n \text{ is an } n \times 1 \text{ matrix)}$$

The transpose of a column vector is a row vector:

$$x^\mathsf{T} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \qquad \text{(i.e. a } 1 \times n \text{ matrix)}$$

Matrix basics

Two vectors $x, y \in \mathbb{R}^n$ can be multiplied together in two ways. Both are valid matrix multiplications:

- **inner product**: produces a scalar.

$$x^\mathsf{T} y = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1 y_1 + \cdots + x_n y_n$$

  Also called "dot product". Often written $x \cdot y$ or $\langle x, y \rangle$.

- **outer product**: produces an $n \times n$ matrix.

$$x y^\mathsf{T} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & \dots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \dots & x_n y_n \end{bmatrix}$$

Matrix basics

- Matrices and vectors can be stacked and combined to form bigger matrices as long as the dimensions agree. e.g. If $x_1, \ldots, x_m \in \mathbb{R}^n$, then $X = \begin{bmatrix} x_1 & x_2 & \ldots & x_m \end{bmatrix} \in \mathbb{R}^{m \times n}$.

- Matrices can also be concatenated in blocks. For example:

$$Y = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

if A, C have same number of columns, A, B have same number of rows, etc.

- Matrix multiplication also works with block matrices!

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} AP + BQ \\ CP + DQ \end{bmatrix}$$

as long as A has as many columns as P has rows, etc.

Linear and affine functions

- A function $f(x_1, \ldots, x_m)$ is *linear* in the variables $x_1, \ldots, x_m$ if there exist constants $a_1, \ldots, a_m$ such that

$$f(x_1, \ldots, x_m) = a_1 x_1 + \cdots + a_m x_m = a^{\mathsf{T}} x$$

- A function $f(x_1, \ldots, x_m)$ is *affine* in the variables $x_1, \ldots, x_m$ if there exist constants $b, a_1, \ldots, a_m$ such that

$$f(x_1, \ldots, x_m) = a_0 + a_1 x_1 + \cdots + a_m x_m = a^{\mathsf{T}} x + b$$

Examples:

1. $3x - y$ is linear in (x, y).
2. $-6x + 7y - 1$ is affine in (x, y).
3. $x^2 + y^2$ is not linear or affine.

- N.B.: Some texts use linear and affine interchangeably

Linear and affine functions

Several linear or affine functions can be combined:

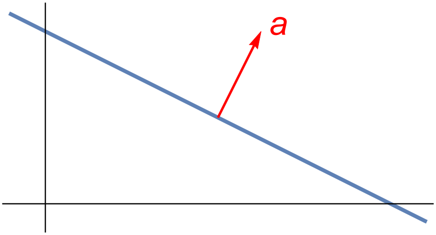
$$\begin{matrix} a_{11}x_1 + \cdots + a_{1n}x_n + b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n + b_2 \\ \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n + b_m \end{matrix} \implies \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

which can be written simply as $Ax + b$. Same definitions apply:

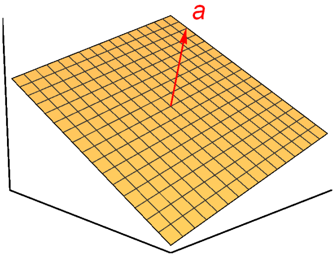
- A vector-valued function $F(x)$ is *linear* in x if there exists a constant matrix A such that $F(x) = Ax$.
- A vector-valued function $F(x)$ is *affine* in x if there exists a constant matrix A and vector b such that $F(x) = Ax + b$.

Geometry of affine equations

- The set of points $x \in \mathbb{R}^n$ that satisfies a linear equation $a_1 x_1 + \cdots + a_n x_n = 0$ (or $a^\mathsf{T} x = 0$) is called a *hyperplane*. The vector a is *normal* to the hyperplane.
- If the right-hand side is nonzero: $a^\mathsf{T} x = b$, the solution set is called an *affine hyperplane*, (it's a shifted hyperplane).



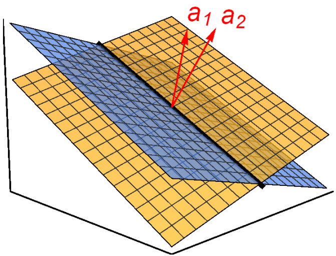
Affine hyperplane in 2D



Affine hyperplane in 3D

Geometry of affine equations

- The set of points $x \in \mathbb{R}^n$ satisfying many linear equations $a_{i1}x_1 + \cdots + a_{in}x_n = 0$ for $i = 1, \dots, m$ (or $Ax = 0$) is called a *subspace* (the intersection of many hyperplanes).
- If the right-hand side is nonzero: $Ax = b$, the solution set is called an *affine subspace*, (it's a shifted subspace).



Intersections of affine hyperplanes are affine subspaces.

Geometry of affine equations

The *dimension* of a subspace is the number of independent directions it contains: the *size of the largest set of linearly independent vectors* in the subspace.

A line has dimension 1, a plane has dimension 2, and so on.

Hyperplanes are subspaces!

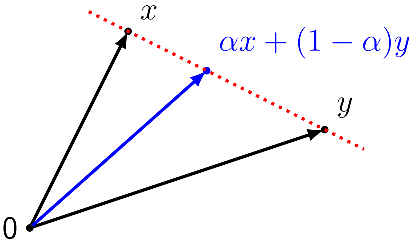
- A hyperplane in $\mathbb{R}^n$ is a subspace of dimension $n-1$.
- The intersection of k hyperplanes has dimension at least $n-k$ ("at least" because of potential redundancy).

Affine combinations

If $x, y \in \mathbb{R}^n$, then the combination

$$w = \alpha x + (1-\alpha)y \quad \text{for some } \alpha \in \mathbb{R}$$

is called an *affine combination*.



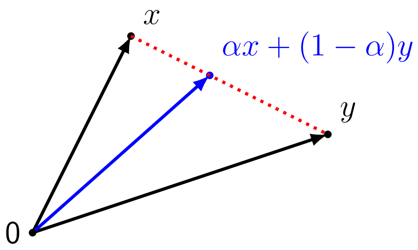
If $Ax = b$ and $Ay = b$, then $Aw = b$. So affine combinations of points in an (affine) subspace also belong to the subspace.

Convex combinations

If $x, y \in \mathbb{R}^n$, then the combination

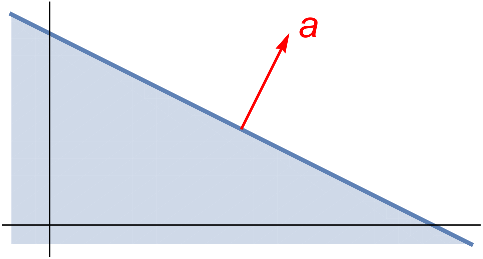
$$w = \alpha x + (1 - \alpha) y \quad \text{for some } 0 \leq \alpha \leq 1$$

is called a *convex combination* (for reasons we will learn later). It's the line segment that connects x and y.



Geometry of affine inequalities

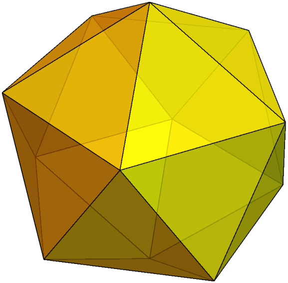
- The set of points $x \in \mathbb{R}^n$ that satisfies a linear inequality $a_1x_1 + \cdots + a_nx_n \leq b$ (or $a^\mathsf{T}x \leq b$) is called a *halfspace*. The vector a is *normal* to the halfspace and b shifts it.
- Define $w = \alpha x + (1-\alpha)y$ where $0 \leq \alpha \leq 1$. If $a^\mathsf{T}x \leq b$ and $a^\mathsf{T}y \leq b$, then $a^\mathsf{T}w \leq b$.



Halfspace

Geometry of affine inequalities

- The set of points $x \in \mathbb{R}^n$ satisfying many linear inequalities $a_{i1}x_1 + \cdots + a_{in}x_n \le b_i$ for $i = 1, \ldots, m$ (or $Ax \le b$) is called a *polyhedron* (the intersection of many halfspaces). Some sources use the term *polytope* instead.
- As before: let $w = \alpha x + (1 - \alpha)y$ where $0 \le \alpha \le 1$. If $Ax \le b$ and $Ay \le b$, then $Aw \le b$.



Intersections of halfspaces are polyhedra.

The linear program

A linear program is an optimization model with:

- real-valued variables ($x \in \mathbb{R}^n$)
- affine objective function ($c^\mathsf{T}x + d$), can be min or max.
- constraints may be:
  - affine equations ($Ax = b$)
  - affine inequalities ($Ax \le b$ or $Ax \ge b$)
  - combinations of the above
- individual variables may have:
  - box constraints ($p_i \le x_i$, or $x_i \le q_i$, or $p_i \le x_i \le q_i$, where p_i and q_i are parameters, not variables)
  - no constraints (x_i is unconstrained)

There are many equivalent ways to express the same LP

Standard form

- Every LP can be put in the form:

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \text{imize} \quad & c^\mathsf{T} x \\ \text{subject to:} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

- We'll call this the *standard form* of a LP.
- (Unfortunately, there are multiple definitions of "standard form" but let's use this one for purposes of this class.)

Back to Top Brass

$$\begin{aligned} \max_{f,s} \quad & 12f + 9s \\ \text{s.t.} \quad & 4f + 2s \leq 4800 \\ & f + s \leq 1750 \\ & 0 \leq f \leq 1000 \\ & 0 \leq s \leq 1500 \end{aligned}$$

$$\Longrightarrow$$

$$\begin{aligned} \max_{f,s} \quad & \begin{bmatrix} 12 \\ 9 \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} f \\ s \end{bmatrix} \\ \text{s.t.} \quad & \begin{bmatrix} 4 & 2 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f \\ s \end{bmatrix} \leq \begin{bmatrix} 4800 \\ 1750 \\ 1000 \\ 1500 \end{bmatrix} \\ & \begin{bmatrix} f \\ s \end{bmatrix} \geq 0 \end{aligned}$$

This is in standard form, with:

$$A = \begin{bmatrix} 4 & 2 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 4800 \\ 1750 \\ 1000 \\ 1500 \end{bmatrix}, \quad c = \begin{bmatrix} 12 \\ 9 \end{bmatrix}, \quad x = \begin{bmatrix} f \\ s \end{bmatrix}$$

Transformation tricks

1. converting min to max or vice versa (take the negative):
$$\min_x f(x) = -\max_x (-f(x))$$

2. reversing inequalities (flip the sign):
$$Ax \le b \quad \iff \quad (-A)x \ge (-b)$$

3. equalities to inequalities (double up):
$$f(x) = 0 \quad \iff \quad f(x) \ge 0 \quad \text{and} \quad f(x) \le 0$$

4. inequalities to equalities (add slack):
$$f(x) \le 0 \quad \iff \quad f(x) + s = 0 \quad \text{and} \quad s \ge 0$$

Transformation tricks

5. unbounded to bounded (add difference):

$$x \in \mathbb{R} \quad \iff \quad u \geq 0, \quad v \geq 0, \quad \text{and} \quad x = u - v$$

6. bounded to unbounded (convert to inequality):

$$p \leq x \leq q \quad \iff \quad \begin{bmatrix} 1 \\ -1 \end{bmatrix} x \leq \begin{bmatrix} q \\ -p \end{bmatrix}$$

7. bounded to nonnegative (shift the variable)

$$p \leq x \leq q \quad \iff \quad 0 \leq (x - p) \quad \text{and} \quad (x - p) \leq (q - p)$$

More complicated example

Convert the following LP to standard form:

$$\begin{aligned} \underset{p,q}{\text{minimize}} \quad & p + q \\ \text{subject to:} \quad & 5p - 3q = 7 \\ & 2p + q \geq 2 \\ & 1 \leq q \leq 4 \end{aligned}$$

notebook: Standard Form.ipynb

More complicated example

Equivalent LP (standard form):

$$
\begin{aligned}
\underset{u,v,w}{\text{maximize}} \quad & -u + v - w \\
\text{subject to:} \quad & -5u + 5v + 3w \le -10 \\
& 5u - 5v - 3w \le 10 \\
& -2u + 2v - w \le -1 \\
& w \le 3 \\
& u, v, w \ge 0
\end{aligned}
$$

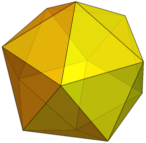
where: $p := u - v, \quad q := w + 1$

and: (original cost) $= -$(new cost) $+ 1$

LPs and polyhedra

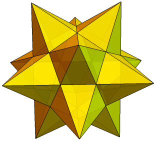
Linear programs have polyhedral feasible sets:

$$\{x \mid Ax \leq b\} \qquad \Longrightarrow$$



Can every polyhedron be expressed as $Ax \leq b$?

Not this one...

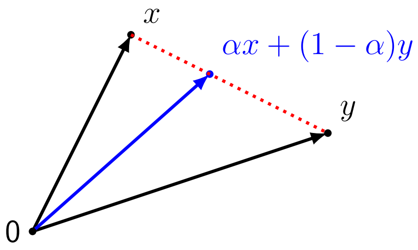


LPs and polyhedra

If $x, y \in \mathbb{R}^n$, then the linear combination

$$w = \alpha x + (1-\alpha) y \quad \text{for some } 0 \leq \alpha \leq 1$$

is called a *convex combination*. As we vary α, it traces out the line segment that connects x and y.



LPs and polyhedra

Note that when we say that $c \le d$ where c and d are two *vectors* of the same dimension, we mean that *every component of c is less than or equal to the corresponding component of d*.

We can have vectors for which neither $c \le d$ nor $c \ge d$ is true!

> If $Ax \le b$ and $Ay \le b$, and w is a convex combination of x and y, then $Aw \le b$.

Proof: Suppose $w = \alpha x + (1 - \alpha)y$.

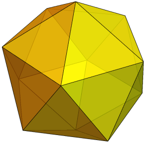
$$\begin{aligned} Aw &= A\left(\alpha x + (1 - \alpha)y\right) \\ &= \alpha Ax + (1 - \alpha)Ay \\ &\le \alpha b + (1 - \alpha)b = b \end{aligned}$$

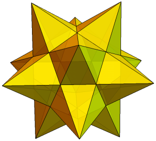
Therefore, $Aw \le b$, which is what we were trying to prove.

LPs and polyhedra

The previous result implies that every polyhedron describable as $Ax \le b$ must contain all convex combinations of its points.

- Such polyhedra are called *convex*.
- Informal definition: if you were to "shrink-wrap" it, the entire polyhedron would be covered with no extra space.

Convex: 

Not convex: 

Goes the other way too: every convex polyhedron can be represented as $Ax \le b$ for appropriately chosen A and b.

Next...

- General modeling
- Cases of LP
- Start working on homework 1!