ORF 522

Linear Programming and Convex Analysis

Canonical & Standard Programs, Linear Equations in an LP Context

Marco Cuturi

Today

- A typology for linear programs.
- Linear equations reminders,
- Basic solutions, the kind of solutions we will be interested in,
- Hyperplanes, or how to visualize linear objectives/constraints.
- A few grams of topology to define halfspaces.

Typology of Linear Programs

Remember...

- the general form of linear programs:

$$
\begin{array}{ll}
\text{max } \textit{or} \text{ min} & z = \boldsymbol{c_1} x_1 + \boldsymbol{c_2} x_2 + \cdots + \boldsymbol{c_n} x_n, \\
\text{subject to} & \left\{ \begin{array}{ccccccccc}
\boldsymbol{a_{11}} x_1 & + & \boldsymbol{a_{12}} x_2 & + & \cdots & + & \boldsymbol{a_{1n}} x_n & \left\{ \begin{array}{c} <,> \\ = \\ \leq,\geq \end{array} \right\} & \boldsymbol{b_1}, \\
\vdots & & & & & & & & \vdots \\
\boldsymbol{a_{m1}} x_1 & + & \boldsymbol{a_{m2}} x_2 & + & \cdots & + & \boldsymbol{a_{mn}} x_n & \left\{ \begin{array}{c} <,> \\ = \\ \leq,\geq \end{array} \right\} & \boldsymbol{b_m},
\end{array} \right. \\
\text{where} & x_1, x_2, \cdots, x_n \geq \mathbf{0}.
\end{array}
$$

- This form is however too vague to be easily usable.
- First step: get rid of the strict inequalities: do not bring much and would only add numerical noise.
- Second step: use matrix and vectorial notations to alleviate.

Notations

Unless explicitly stated otherwise,

- A, B etc... are matrices whose size is clear from context.
- $\mathbf{x}, \mathbf{b}, \mathbf{a}$ are vectors. $\mathbf{a}_1, \mathbf{a}_k$ are members of a vector family.
- $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ with vector coordinates x_i in **R**.
- $\mathbf{x} \geq 0$ is meant coordinate-wise, that is $x_i \geq 0$ for $1 \geq i \leq n$
- $\mathbf{x} \neq \mathbf{0}$ means that $\mathbf{x}$ is not the zero vector, i.e. there exists at least one index i such that $x_i \neq 0$.
- $\mathbf{x}^T$ is the transpose $\begin{bmatrix} x_1, \cdots, x_n \end{bmatrix}$ of $\mathbf{x}$.

Linear Program

Common representation for all these programs?

- Would help in developing both theory & algorithms.
- Also helps when developing software, solvers, etc

The answer is yes. . .

- There are 2: **standard form** and **canonical form**

Terminology

- A linear program in **canonical** form is the program

$$\begin{array}{ll} \text{max or min} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \leq \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{array}$$

$\mathbf{b} \geq 0 \Rightarrow$ **feasible canonical form** (just a convention)

- A linear program in **standard** form is the program

$$\text{max or min} \quad \mathbf{c}^T\mathbf{x} \tag{1}$$

$$\text{subject to} \quad A\mathbf{x} = \mathbf{b}, \tag{2}$$

$$\mathbf{x} \geq \mathbf{0}. \tag{3}$$

Linear Programs: a look at the canonical form

Canonical form linear program

- **Maximize** the objective
- Only **inequality** constraints
- All variables should be **positive**

Example:

$$\begin{array}{lrcrcrcr} \text{maximize} & 5x_1 & + & 4x_2 & + & 3x_3 & & \\ \text{subject to} & 2x_1 & + & 3x_2 & + & x_3 & \leq & 5 \\ & 4x_1 & + & x_2 & + & 2x_3 & \leq & 11 \\ & 3x_1 & + & 4x_2 & + & 2x_3 & \leq & 8 \\ & & & & x_1, & x_2, \; x_3 & \geq & 0. \end{array}$$

Linear Programs: canonical form

Although more intuitive than the standard form, the canonical is not the most useful,

- We will formulate the simplex method on problems with **equality constraints**, that is **standard forms**.
- Solvers do not all agree on this input format. MATLAB for example uses:

$$
\begin{array}{ll}
\text{minimize} & \sum_i c_i x_i \\
\text{subject to} & \sum_{j=1}^{n} A_{ij} x_j \leq b_i, \quad i = 1, \ldots, m_1 \\
& \sum_{j=1}^{n} B_{ij} x_j = d_i, \quad i = 1, \ldots, m_2 \\
& l_i \leq x_i \leq u_i, \quad i = 1, \ldots, n
\end{array}
$$

- Ultimately: this is a **non-issue**, we can easily switch from one form to the other. . .

Linear Programs: standard & canonical form

equalities $\Rightarrow$ inequalities

- What if the original problem has equality constraints?
- Replace equality constraints by two inequality constraints.
- The inequality

$$2x_1 + 3x_2 + x_3 = 5,$$

is equivalent to

$$2x_1 + 3x_2 + x_3 \leq 5 \quad \text{and} \quad 2x_1 + 3x_2 + x_3 \geq 5$$

- The new problem is **equivalent** to the previous one. . .

Linear Programs: standard & canonical form

inequalities $\Rightarrow$ equalities

- The opposite direction works too. . .
- Turn inequality constraints into equality constraints by **adding variables**.
- The inequality

$$2x_1 + 3x_2 + x_3 \leq 5,$$

  is equivalent to

$$2x_1 + 3x_2 + x_3 + w_1 = 5 \quad \text{and} \quad w_1 \geq 0,$$

- The new variable is called a **slack** variable (one for each inequality in the program). . .
- The new problem is **equivalent** to the previous one. . .

Linear Programs: standard & canonical form

free variable $\Rightarrow$ positive variables

- What about free variables?
- A free variable is simply the difference of its positive and negative parts. Again the solution is again **adding variables**.
- If the variable y is free, we can write it

$$y_1 = y_2 - y_3 \quad \text{and} \quad y_2, y_3 \geq 0,$$

- We add two positive variables for each free variable in the program.
- Again, the new problem is **equivalent** to the previous one.

Linear Programs: standard & canonical form

minimizing $\Rightarrow$ maximizing

- What happens when the objective is to minimize? We can use the fact that

$$\min_x f(x) = -\max_x -f(x)$$

- In a linear program this means

$$\text{minimize} \quad 6x_1 \quad - \quad 3x_2 \quad + \quad 5x_3$$

  becomes:

$$- \text{ maximize} \quad -6x_1 \quad + \quad 3x_2 \quad - \quad 5x_3$$

That's all we need to convert all linear programs in standard form. . .

Linear Programs: standard & canonical form

Example. . .

$$\begin{array}{lrcrcrcr} \text{minimize} & 2x_1 & - & 4x_2 & + & x_3 & & \\ \text{subject to} & 2x_1 & + & 7x_2 & + & x_3 & = & 5 \\ & 4x_1 & + & x_2 & + & 9x_3 & \leq & 11 \\ & 3x_1 & + & 4x_2 & + & 2x_3 & = & 8 \\ & & & & & x_1,\ x_2 & \geq & 0. \end{array}$$

This program has one free variable (x_3) and one inequality constraint. It's a minimization problem. . .

Linear Programs: standard & canonical form

We first turn it into a **maximization**. . .

$$
\begin{array}{lrcrcrcr}
-\ \text{maximize} & -2x_1 & + & 4x_2 & - & x_3 & & \\
\text{subject to} & 2x_1 & + & 7x_2 & + & x_3 & = & 5 \\
 & 4x_1 & + & x_2 & + & 9x_3 & \leq & 11 \\
 & 3x_1 & + & 4x_2 & + & 2x_3 & = & 8 \\
 & & & & & x_1,\ x_2 & \geq & 0.
\end{array}
$$

Just switch the signs in the objective. . .

Linear Programs: standard & canonical form

We then turn the inequality into an **equality** constraint by adding a slack variable. . .

$$
\begin{array}{lrcrcrcrcr}
-\text{ maximize} & -2x_1 & + & 4x_2 & - & x_3 & & & & \\
\text{subject to} & 2x_1 & + & 7x_2 & + & x_3 & & & = & 5 \\
& 4x_1 & + & x_2 & + & 9x_3 & + & w_1 & = & 11 \\
& 3x_1 & + & 4x_2 & + & 2x_3 & & & = & 8 \\
& & & & & x_1,\ x_2 & & & \geq & 0.
\end{array}
$$

Now, we only need to get rid of the free variable. . .

Linear Programs: standard & canonical form

We replace the free variable by a difference of two **positive** ones:

$$
\begin{array}{lrcrcrcrcr}
-\text{ maximize} & -2x_1 & + & 4x_2 & - & (x_4 - x_5) & & & & \\
\text{subject to} & 2x_1 & + & 7x_2 & + & x_4 - x_5 & & & = & 5 \\
 & 4x_1 & + & x_2 & + & 9x_4 - 9x_5 & + & w_1 & = & 11 \\
 & 3x_1 & + & 4x_2 & + & 2x_4 - 2x_5 & & & = & 8 \\
 & & & & & x_1,\ x_2,\ x_4,\ x_5 & & & \geq & 0.
\end{array}
$$

- That's it, we've reached a standard form.
- The simplex algorithm is easier to write with this form.

To sum up...

- A linear program in **standard** form is the program

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} = \mathbf{b}, \\ & \mathbf{x} \geq \mathbf{0}. \end{array} \tag{4}$$

  where

  - $\mathbf{c}, \mathbf{x} \in \mathbf{R}^n$ – the objective,
  - $A \in \mathbf{R}^{m \times n}$ and $\mathbf{b} \in \mathbf{R}^m$ – the equality constraints,
  - $\mathbf{x} \geq \mathbf{0}$ means that for $\mathbf{x} = (x_1, \ldots, x_n), x_i \geq 0$ for $1 \leq i \leq n$.

- From now on we focus on

  - **linear constraints** $A\mathbf{x} = \mathbf{b}$,
  - **objective function** $\mathbf{c}^T\mathbf{x}$,

  separately.

- $\mathbf{x} \geq \mathbf{0}$ will reappear when we study convexity.

Linear Equations

Linear Equations

The usual linear equations we know, $m = n$

- In the usual linear algebra setting, A is square of size n and invertible.
- Straightforward: $\{\mathbf{x} \in \mathbf{R}^n | A\mathbf{x} = \mathbf{b}\}$ **is a singleton,** $\{A^{-1}\mathbf{b}\}$.
- Focus: find **efficiently** that **unique** solution. Many methods (Gaussian pivot etc.)

In classic statistics, most often $m \gg n$

- A few explicative variables, a lot of observations.
- Generally $\{\mathbf{x} \in \mathbf{R}^n | A\mathbf{x} = \mathbf{b}\} = \emptyset$ so we need to tweak the problem
- Least-squares regression: select $\mathbf{x}_0 \mid \mathbf{x}_0 = \text{argmin} \, |A\mathbf{x} - \mathbf{b}|^2$
- More advanced, penalized LS regression: $\mathbf{x}_0 = \text{argmin}(|A\mathbf{x} - \mathbf{b}|^2 + \lambda \|\mathbf{x}\|)$

Linear Equations

On the other hand, in an LP setting where usually $m < n$

- $\{\mathbf{x} \in \mathbf{R}^n | A\mathbf{x} = \mathbf{b}\}$ is a wider set of candidates, a convex set.
- In LP, a linear criterion is used to choose one of them.
- In other fields, such as **compressed sensing**, other criterions are used.
- Today we start studying some simple properties of the set $\{\mathbf{x} \in \mathbf{R}^n | A\mathbf{x} = \mathbf{b}\}$.

Linear Equations

- **Linear Equation**: $A\mathbf{x} = \mathbf{b}$, m equations.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array}$$

- Writing $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$ we have n columns $\in \mathbf{R}^m$.
- Add now $\mathbf{b}$: $A_b = [A, b] \in \mathbf{R}_{m\times n+1}$.
- remember: a solution to $A\mathbf{x} = \mathbf{b}$ is a vector $\mathbf{x}$ such that

$$\sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b},$$

that is the $\mathbf{b}$ and $\mathbf{a}$'s should be **linearly dependent** (l.d.) for everything to work.

Linear Equations

Two cases (note that $\mathbf{Rank}(A)$ cannot be $> \mathbf{Rank}(A_b)$)

- (i) $\mathbf{Rank}(A) < \mathbf{Rank}(A_b)$; $\mathbf{b}$ and $\mathbf{a}$'s are **linearly independent** (l.i.). *no solution*.
- (ii) $\mathbf{Rank}(A) = \mathbf{Rank}(A_b) = k$; every column of A_b, $\mathbf{b}$ in particular, can be expressed as a linear combination of k other columns of the matrix $a_{i_1}, \cdots, a_{i_k}$. Namely, $\exists \mathbf{x}$ such that

$$\sum_{j=1}^{k} x_{i_j} \mathbf{a}_{i_j} = \mathbf{b}.$$

In practice

- if $m = n = k$, then there is a unique solution: $\mathbf{x} = A^{-1}\mathbf{b}$;
- Usually $\mathbf{Rank}(A) = k \leq m < n$ and we have a plenty of solutions;
- We assume from now on that $\mathbf{Rank}(A) = \mathbf{Rank}(A_b) = m$.

Linear Equation Solutions

- if $\mathbf{x}_1$ and $\mathbf{x}_2$ are two different solutions, then $\forall \lambda \in \mathbf{R}, \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ is a solution.
- $\mathbf{Rank}(A) = m$. There are m independent columns. Suppose we reorder them so that $\mathbf{a}_1, \cdots, \mathbf{a}_m$ are linearly independent.
- Then

$$A = \left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1m} & a_{1m+1} & a_{1m+2} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2m} & a_{2m+1} & a_{2m+2} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} & a_{mm+1} & a_{mm+2} & \cdots & a_{mn} \end{array}\right] = [B, R]$$

- B is $m \times m$ square, R is $m \times (n-m)$ rectangular.

Linear Equation Solutions

- suppose we divide $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_\beta \end{bmatrix}$ where $\mathbf{x}_B \in \mathbf{R}^m$ and $\mathbf{x}_\beta \in \mathbf{R}^{m-n}$
- If $A\mathbf{x} = \mathbf{b}$ then $B\mathbf{x}_B + R\mathbf{x}_\beta = \mathbf{b}$. Since B is non-singular, we have

$$\mathbf{x}_B = B^{-1}(\mathbf{b} - R\mathbf{x}_\beta),$$

  which shows that we can assign **arbitrary** values to $\mathbf{x}_\beta$ and obtain different points $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$.
- Solutions are parameterized by $\mathbf{x}_\beta$... a bit problematic since R is the "discarded" part.
- We choose $\mathbf{x}_\beta = \mathbf{0}$ and focus on the choice of B.

Basic Solutions

Basic Solutions

Definition 1. *Consider* $A\mathbf{x} = \mathbf{b}$ *and suppose* $\mathbf{Rank}(A) = m < n$. *Let* $\mathbf{I} = (i_1, \cdots, i_m)$ *be a list of indexes corresponding to* m ***linearly independent*** *columns taken among the* n *columns of* A.

- *We call the* m *variables* $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \cdots, \mathbf{x}_{i_m}$ *of* $\mathbf{x}$ *its* ***basic variables***,
- *the other variables are called* ***non-basic***.

If $\mathbf{x}$ *is a vector such that* $A\mathbf{x} = \mathbf{b}$ *and all* ***its non-basic variables are equal to 0*** *then* $\mathbf{x}$ *is a basic solution.*

Basic Solutions

- When reordering variables as in the previous slide, and defining $B = [\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}]$ we can set $\mathbf{x}_\beta = \mathbf{0}$. Then $\mathbf{x}_B = B^{-1}\mathbf{b}$ and

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{0} \end{bmatrix},$$

  and we have a **basic solution**.

- Sidenote: a **basic feasible solution** to an LP Equation (4) is such that $\mathbf{x}$ is basic **and** $\mathbf{x} \geq 0$.

Basic Solutions

- More generally, let

$$B_{\mathbf{I}} = [\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}],$$

$$R_{\mathbf{O}} = [\mathbf{a}_{o_1}, \cdots, \mathbf{a}_{o_{m-n}}],$$

  where $\mathbf{O} = \{1, \cdots, n\} \setminus \mathbf{I} = (o_1, \cdots, o_{m-n})$ is the complementary of $\mathbf{I}$ in $\{1, \cdots, n\}$ in increasing order.

- $\mathbf{I}$ contains the indexes of vectors **in** the basis, $\mathbf{O}$ contains the indexes of vectors **outside** the basis.

- Equivalently set $\mathbf{x}_{\mathbf{I}} = \begin{bmatrix} x_{i_1} \\ \vdots \\ x_{i_m} \end{bmatrix}, \mathbf{x}_{\mathbf{O}} = \begin{bmatrix} x_{o_1} \\ \vdots \\ x_{o_{n-m}} \end{bmatrix}.$

- $A\mathbf{x} = B_{\mathbf{I}}\mathbf{x}_{\mathbf{I}} + R_{\mathbf{O}}\mathbf{x}_{\mathbf{O}}$

Basic Solutions

The two things to remember so far:

- **A list $\mathbf{I}$ of m independent columns** $\leftrightarrow$ **One basic solution $\mathbf{x}$, with $\mathbf{x_I} = B_{\mathbf{I}}^{-1}\mathbf{b}$ and $\mathbf{x_O} = \mathbf{0}$**
- We are **not** interested in **all** basic solutions, only a subset: **basic feasible solutions**.

Basic Solutions: Degeneracy

Definition 2. *A* ***basic*** *solution to* $A\mathbf{x} = \mathbf{b}$ *is* ***degenerate*** *if one or more of the* m ***basic*** *variables is equal to zero.*

- For a **basic solution**, $\mathbf{x_O}$ is always $\mathbf{0}$. On the other hand, we do not expect elements of $\mathbf{x_I}$ to be zero.
- This is **degeneracy** which appears whenever there is one or more components of $\mathbf{x_I}$ which are zero.

Basic Solutions: Example

- Consider $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 3 \\ 0 & 1 & 2 & 1 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

We start by choosing $\mathbf{I}$:

- $\mathbf{I} = (1,2)$. $B_\mathbf{I} = [\mathbf{a}_1, \mathbf{a}_2] = \left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right] \rightarrow \mathbf{x}_\mathbf{I} = \left[\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right]; \mathbf{x} = \left[\begin{smallmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{smallmatrix}\right]$ is **basic**.

- $\mathbf{I} = (1,4)$. $B_\mathbf{I} = [\mathbf{a}_1, \mathbf{a}_4] = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right] \rightarrow \mathbf{x}_\mathbf{I} = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]; \mathbf{x} = \left[\begin{smallmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{smallmatrix}\right]$ is **basic**.

- $\mathbf{I} = (2,5)$. $B_\mathbf{I} = [\mathbf{a}_2, \mathbf{a}_5] = \left[\begin{smallmatrix} 1 & 3 \\ 0 & 3 \end{smallmatrix}\right] \rightarrow \mathbf{x}_\mathbf{I} = \left[\begin{smallmatrix} 0 \\ \frac{1}{3} \end{smallmatrix}\right]; \mathbf{x} = \left[\begin{smallmatrix} 0 \\ \mathbf{0} \\ 0 \\ 0 \\ \frac{1}{3} \end{smallmatrix}\right]$ is **degenerate basic**

note that $\mathbf{a}_5$ and $\mathbf{b}$ are colinear...

Non-degeneracy

Theorem 1. *A necessary and sufficient condition for the existence and non-degeneracy of all basic solutions of* $A\mathbf{x} = \mathbf{b}$ *is the* ***linear independence*** *of* ***every set of*** m ***columns of*** A_b*, the augmented matrix.*

Proof. • **Proof strategy**: $\Rightarrow$ the existence of all possible basic solutions is already a good sign: all families of m columns of A are l.i. What we need is show that $m-1$ columns of A plus $\mathbf{b}$ are also l.i.

• $\Leftarrow$ if all m columns choices are independent, basic solutions exist, and are non-degenerate because $\mathbf{b}$ is l.i. with any combination of $m-1$ columns.

■

Non-degeneracy

Proof. • $\Rightarrow$: Let $I = (i_1, \cdots, i_m)$ a family of indexes.

- The basic solution associated with I exists and is non-degenerate. $\mathbf{b} \neq 0$
- Hence by definition $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}\}$ is l.i. and $\mathbf{b} = \sum_{k=1}^m x_k \mathbf{a}_{i_k}$.
- For a given r, suppose $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots, \mathbf{a}_{i_m}, \mathbf{b}\}$ is l.d.
- Then $\exists(\alpha_1, \cdots, \alpha_{r-1}, \alpha_{r+1}, \alpha_m)$ and β such that

$$\beta \mathbf{b} + \sum_{k=1, k\neq r}^{m} \alpha_k \mathbf{a}_{i_k} = \mathbf{0}.$$

  Note that necessarily $\beta \neq 0$ (otherwise $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots, \mathbf{a}_{i_m}\}$ is l.d)
- Contradiction: degenerate solution for I, $(-\frac{\alpha_1}{\beta}, \cdots, -\frac{\alpha_{r-1}}{\beta}, 0, -\frac{\alpha_{r+1}}{\beta}, -\frac{\alpha_m}{\beta})$

• $\Leftarrow$: Let $I = (i_1, \cdots, i_m)$ a family of indexes.

- A basic solution exists, $\sum_{k=1}^m x_k \mathbf{a}_{i_k} = \mathbf{b}$
- Suppose it is degenerate, i.e. $x_r = 0$. Then $\sum_{k=1, k\neq r}^m x_k \mathbf{a}_{i_k} - \mathbf{b} = \mathbf{0}$
- Contradiction: $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots, \mathbf{a}_{i_m}, \mathbf{b}\}$, of size m, is l.d.

■

Non-degeneracy

Corollary 1. *Given a basic solution to* $A\mathbf{x} = \mathbf{b}$ *with basic variables* $x_{i_1}, \cdots, x_{i_m}$, *a necessary and sufficient condition for the solution to be non-degenerate is the l.i. of* $\mathbf{b}$ *with every subset of* $m-1$ *columns of* $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}\}$

- In our previous example,

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 3 \\ 0 & 1 & 2 & 1 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; m = 2.$$

- Hence if $\mathbf{I} = (2, 5)$, $[\mathbf{b}, \mathbf{a}_2]$ and $[\mathbf{b}, \mathbf{a}_5]$ should be of rank 2 for the solution not to be degenerate. Yet $[\mathbf{b}, \mathbf{a}_5] = \left[\begin{smallmatrix} 1 & 3 \\ 1 & 3 \end{smallmatrix}\right]$ is clearly of rank 1.

Hyperplanes

Hyperplane

Definition 3. *A hyperplane in* $\mathbf{R}^m$ *is defined by a vector* $\mathbf{c} \neq \mathbf{0} \in \mathbf{R}^m$ *and a scalar* $z \in \mathbf{R}$ *as the set* $\{\mathbf{x} \in \mathbf{R}^m | \mathbf{c}^T\mathbf{x} = z\}$.

$z = 0$**,**

- A hyperplane $H_{\mathbf{c},z}$ contains $\mathbf{0}$ iff $z = 0$.
- In that case $H_{\mathbf{c},0}$ is a **vector subspace** and $\dim(H_{\mathbf{c},0}) = n - 1$

$z \neq 0$**,**

- For $\mathbf{x}_1, \mathbf{x}_2$ easy to check that $\mathbf{c}^T(\mathbf{x}_1 - \mathbf{x}_2) = 0$. In other words $\mathbf{c}$ is orthogonal to vectors lying in the hyperplane.
- $\mathbf{c}$ is called the **normal** of the hyperplane

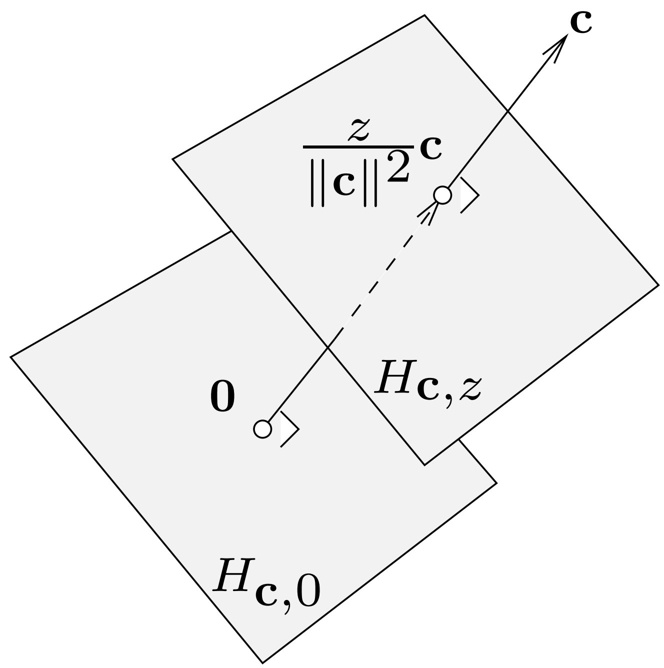
Affine Subspace

Definition 4. *Let V be a vector space and let L be a vector subspace of V. Then given $\mathbf{x} \in V$, the translation $T = L + \mathbf{x} = \{\mathbf{u} + \mathbf{x}, \mathbf{u} \in L\}$ is called an affine subspace of V.*

- the **dimension** of T is the dimension of L.
- T is **parallel** to L.

Affine Hyperplane

- For $\mathbf{c} \neq \mathbf{0}$, $H_{\mathbf{c},0}$ is a Vector subspace of $\mathbf{R}^m$ of dimension $n-1$.
- When $z \neq 0$, $H_{\mathbf{c},z}$ is an affine **hyperplane**: it's easy to see that $H_{\mathbf{c},z} = H_{\mathbf{c},0} + \frac{z}{\|\mathbf{c}\|^2}\mathbf{c}$



Some grams of Topology and Halfspaces

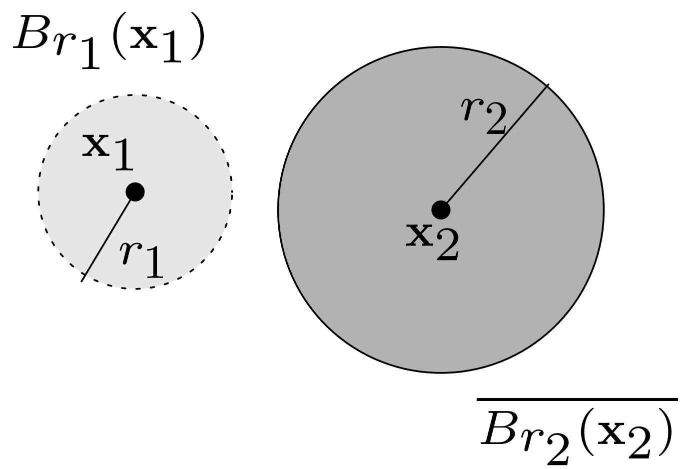
A bit of topology: open and closed balls

- The n dimensional open ball centered at $\mathbf{x}_0$ with radius r is defined as

$$B_r(\mathbf{x}_0) = \{x \in \mathbf{R}^n \text{s.t.} |\mathbf{x} - \mathbf{x}_0| < r\},$$

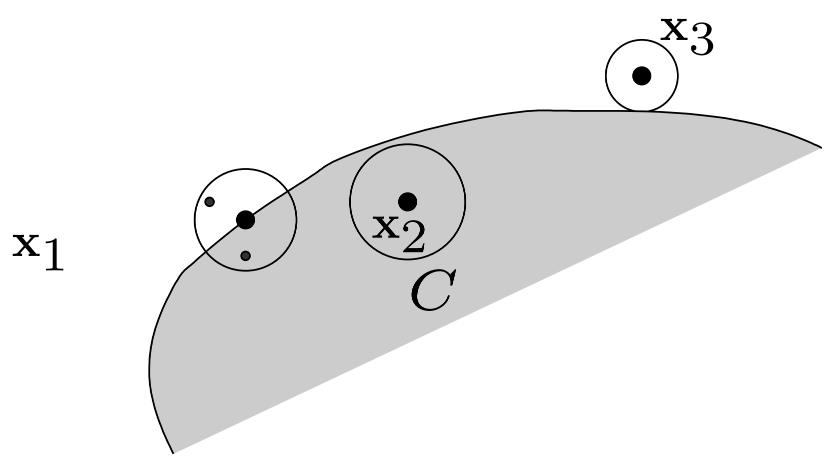
- its closure

$$\overline{B_r(\mathbf{x}_0)} = \{x \in \mathbf{R}^n \text{s.t.} |\mathbf{x} - \mathbf{x}_0| \leq r\},$$



A bit of topology: boundary

- Let $S \subset \mathbf{R}^n$. A point $\mathbf{x}$ is a **boundary point** of S if every open ball centered at $\mathbf{x}$ contains both a point in S and a point in $\mathbf{R}^n \setminus S$.
- A boundary point can either be in S or not in S.



- x_1 is a boundary point, x_2 and x_3 are not.

A bit of topology: open and closed sets

- The set of all boundary points of S is the **boundary** ∂S of S.
- A set is **closed** if $\partial S \subset S$. A set is *open* if $\mathbf{R}^n \setminus S$ is closed.
- Note that there are sets that are **neither** open nor close.
- The **closure** $\overline{S}$ of a set S is $S \cup \partial S$
- The **interior** S^o of a set S is $S \setminus \partial S$
- A set S is *closed* iff $S = \overline{S}$ and *open* iff $S = S^o$.

Halfspaces

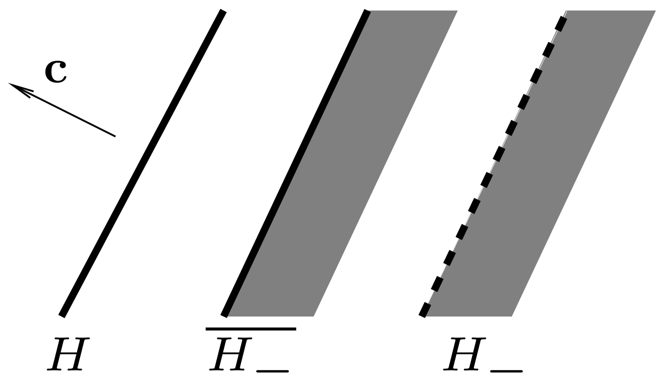
- For a hyperplane H, its complement in $\mathbf{R}^n$ is the union of two sets called **open halfspaces**;

$$\mathbf{R}^n \setminus H = H_+ \cup H_-$$

where

$$\begin{aligned} H_+ &= \{\mathbf{x} \in \mathbf{R}^m | \mathbf{c}^T\mathbf{x} > z\} \\ H_- &= \{\mathbf{x} \in \mathbf{R}^m | \mathbf{c}^T\mathbf{x} < z\} \end{aligned}$$

- $\overline{H_+} = H_+ \cup H$ and $\overline{H_-} = H_- \cup H$ are **closed** halfspaces.



Coming Next

- some basic convexity,
- important interplay between convex sets and hyperplanes,
- starting with some nice results, Caratheodory theorem,
- laying out theoretical fundations to attack the simplex.