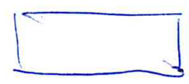
From last time $A = m\overset{n}{[\vec{a}_1, \cdots \vec{a}_n]}$ $\quad m < n$

fat matrix 

rank $(A)$ = max # of linearly independent columns in $A$

linear dependent columns = redundant column

$$A\vec{x} = \vec{b} \qquad \vec{x} = A^{-1}b$$

not square

DL

$$A\vec{x}_1 = b \,, \quad A\vec{x}_2 = \vec{b}$$

$$\vec{x}_1 \neq \vec{x}_2$$

$$\Downarrow$$

$$A\big(\overbrace{\lambda \vec{x}_1 + (1-\lambda)\vec{x}_2}^{\text{solution too}}\big) \qquad \lambda \in \mathbb{R}$$

$$= \lambda \vec{b} + (1-\lambda)\vec{b} = \vec{b}$$

$$\Downarrow$$

$\infty$ <u>many solutions</u>

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 40 & 60 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x \\ y \\ s \\ t \\ u \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 50 \\ 2400 \\ 30 \end{bmatrix}}_{\vec{b}}$$

$x=0$
$s=0$
$t=0$
$y=0$
$u=0$

$s=10,\ t=20$

$$\begin{bmatrix} 1 & 1 & 0 \\ 40 & 60 & 0 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ u \end{bmatrix} \cancel{=} -\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \cancel{s}^{10} \\ \cancel{t}_{20} \end{bmatrix} + \begin{bmatrix} 50 \\ 2400 \\ 300 \end{bmatrix}$$

$s=0=t$
$\Downarrow$
solve for
$x\ y\ u$

$$\begin{array}{c} \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 40 & 60 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 & 0 \end{bmatrix} \\ \begin{matrix} x & y & u & s & t \end{matrix} \end{array} \begin{bmatrix} x \\ y \\ u \\ s \\ t \end{bmatrix} = \begin{bmatrix} 50 \\ 2400 \\ 300 \end{bmatrix}$$

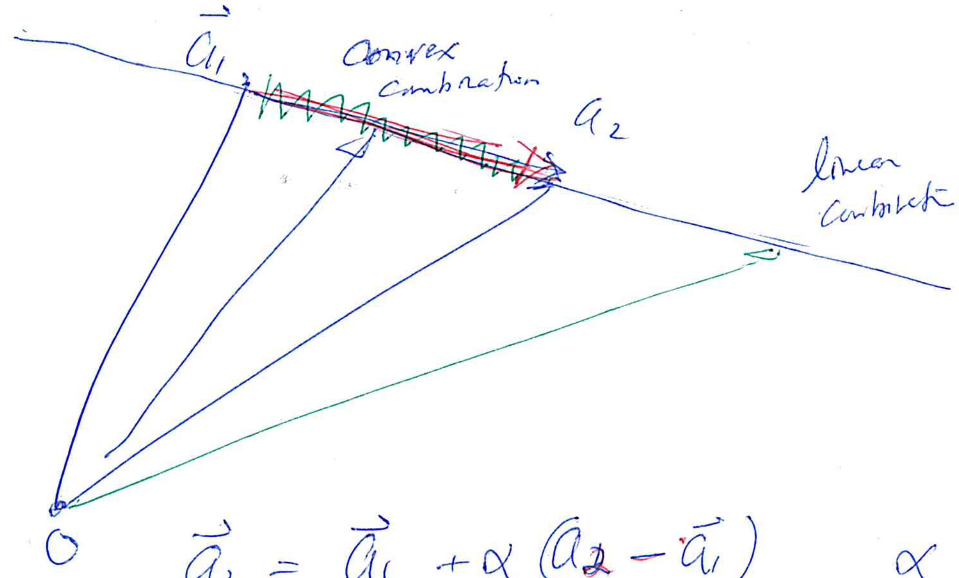
$$\vec{x} = \begin{bmatrix} x \\ y \\ u \\ s \\ t \end{bmatrix} = \begin{bmatrix} * \\ * \\ * \\ 0 \\ 0 \end{bmatrix}$$

$x, y, u$ basic variables
$s\ t$ non-basic variables

Chance $x, y, u$ is zero

$$m\left\{\underbrace{A}_{n}\right.\vec{x} = \vec{b}$$

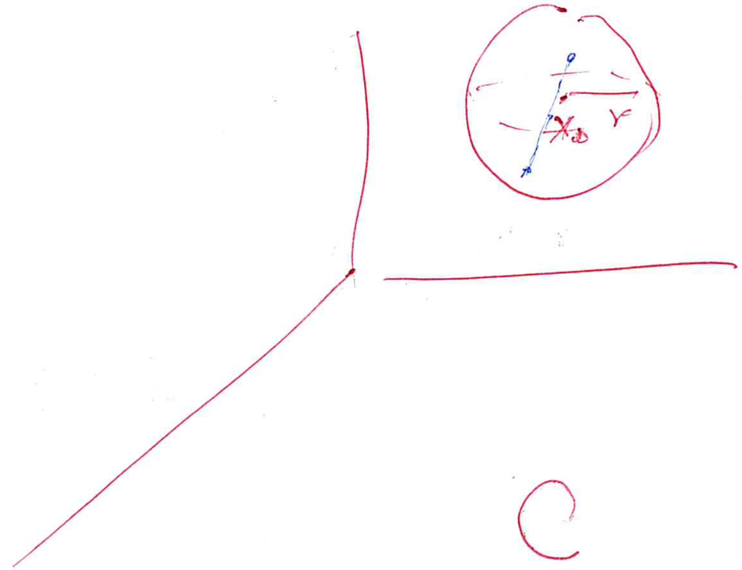
$$\#\text{ "Basic Solutions" } = C^n_m = C^n_{n-m}$$



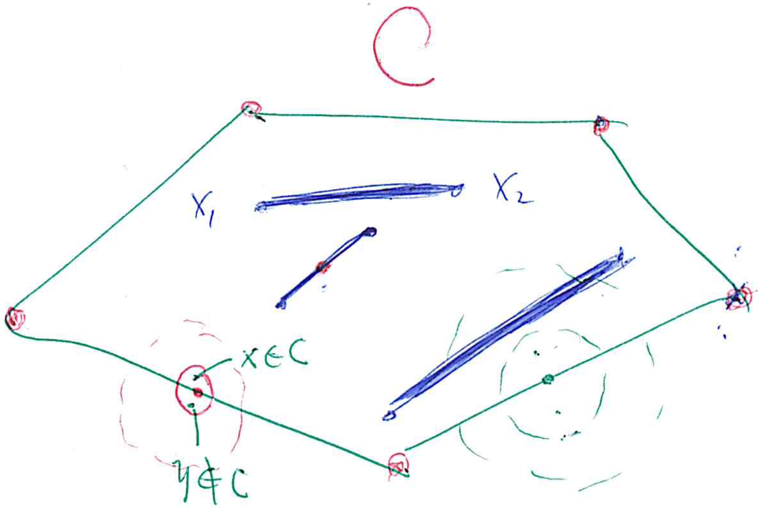

$\vec{a}_3 = \vec{a}_1 + \alpha(\vec{a}_2 - \vec{a}_1)$ $\qquad \alpha \in \mathbb{R}$

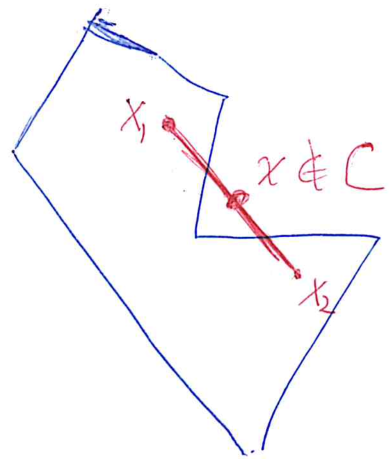
$\vec{a}_3 = (1-\alpha)\vec{a}_1 + \alpha \vec{a}_2$ $\qquad \alpha \in \mathbb{R}$. linear combination

$\vec{a}_4 = (1-\alpha)\vec{a}_1 + \alpha \vec{a}_2$ $\qquad \alpha \in [0,1]$ convex combination

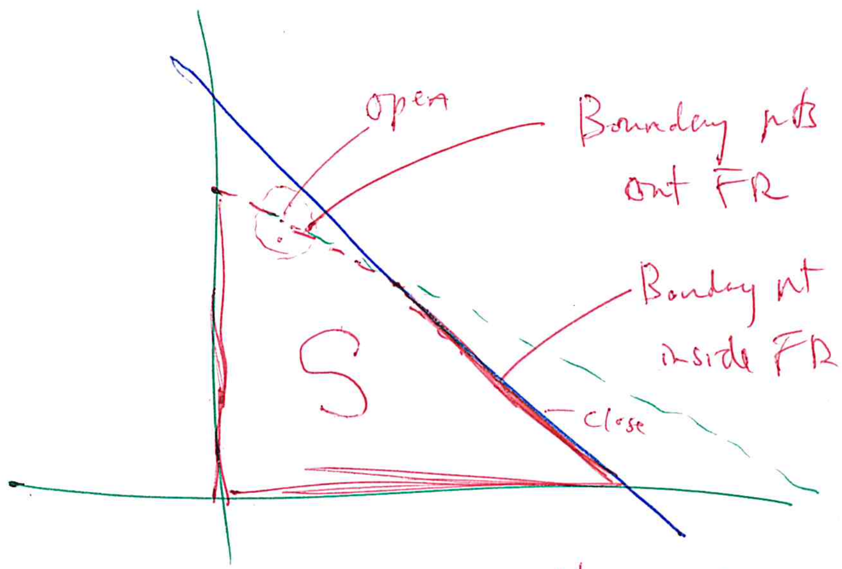



Convex



$$\begin{cases} x + y \le 50 \\ 40x + 60y < 2400 \\ x, y \ge 0 \end{cases}$$




$S \cup \partial S = \overline{S}$

is closed

closure of S

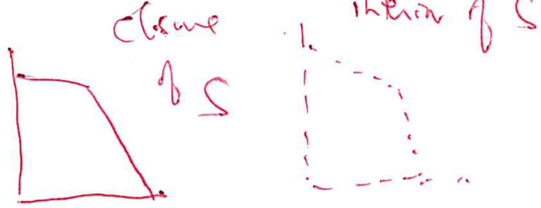
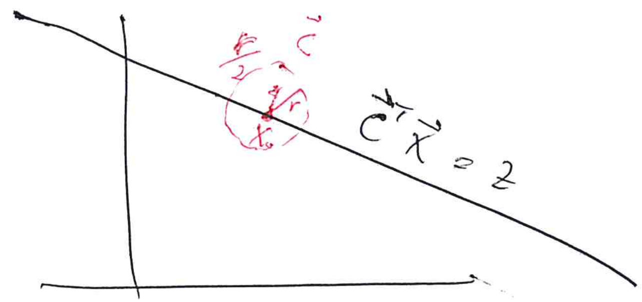


Fig 1.5 $\quad P = \{ x \mid \vec{c}^T x = z \}$ is closed

normal $|\vec{c}| = 1$

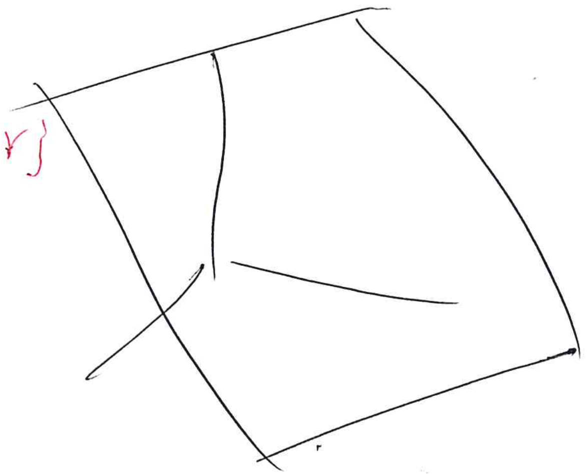
$\boxed{\text{We have to prove this } \partial P \subseteq P}$




$P \subseteq \partial P$ In fact we can prove that $P = \partial P$

$\forall \vec{x}_0 \in P \ \forall r > 0, \ B_r(\vec{x}_0) = \{ \vec{x} \mid |\vec{x} - \vec{x}_0| < r \}$

$B_r(\vec{x}_0) \ni \vec{x}_0$ is inside $P$.



claim

$B_r(\vec{x}_0) \ni \vec{x} = \vec{x}_0 + \frac{r}{2} \vec{c} \notin P$

$\vec{c}^T \vec{x} = \vec{c}^T \vec{x}_0 + \frac{r}{2} \vec{c}^T c = z_0 + \frac{r}{2} |c| = z_0 + \frac{r}{2} \cancel{z_0}$

$\therefore \ \vec{x} \notin P$

$|\vec{x} - \vec{x}_0| = |\frac{r}{2} \vec{c}| = \frac{r}{2} < r. \Rightarrow \vec{x} \in B_r(\vec{x}_0)$

# The Simplex Method for a Two-variable Problem

## 0.1 Interpretation of the Graphical Method

To introduce the basic ideas of the simplex method, we will use an example with only two decision variables $x$ and $y$. We can then see how both the graphical method and the simplex method works. Consider

$$\text{Max} \quad f(x,y) = 30x + 20y \; =$$

$$\text{subject to} \begin{cases} x + \quad y \leq \; 50 \\ 40x + 60y \leq 2400 \\ x, \qquad y \geq \; 0 \end{cases} \tag{0.1}$$

$ax + by \leq c$

$ax + by + u = c$

The graph of the feasible region is given in Figure 0.1.

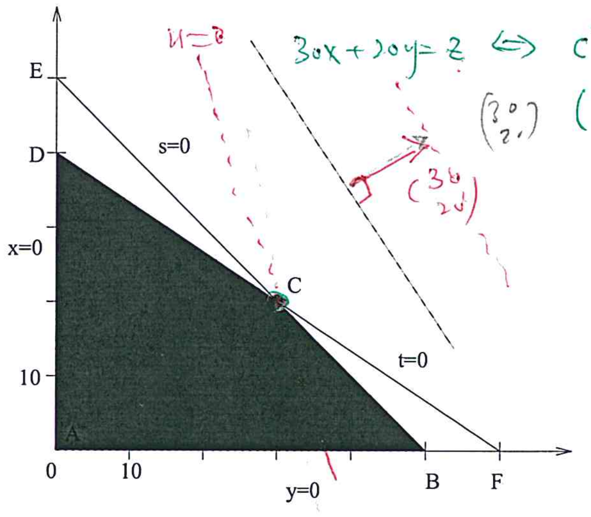



**Figure 0.1.** Feasible region for (0.1)

$30x + 20y = z \iff c^T \vec{x} = z$

$(30, 20)\begin{pmatrix} x \\ y \end{pmatrix} = z$

$\begin{pmatrix} x \neq 0 \\ y \neq 0 \\ s = 0 \\ t = 0 \\ u = 0 \end{pmatrix}$ $\quad \underbrace{\left( \frac{30}{\sqrt{30^2+20^2}} \; \frac{20}{\sqrt{30^2+20^2}} \right)}_{\text{normal of norm } 1}\vec{x}$

$s = 0 \quad t = 0$

$\begin{pmatrix} x \\ y \\ u \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \end{pmatrix} \quad 10^{-13}$

We note that the constraints are inequalities. Since inequalities are difficult to be handled by matrices, we first change them into equalities by adding two more variables

$t = 0, \; u = 0$

$\begin{pmatrix} x \\ y \\ s \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \end{pmatrix}$

Corner

**Definition 1.6.** A point $\mathbf{x}$ is an *extreme point* of a convex set $C$ if there exist no two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2 \in C$ such that $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for some $\lambda \in (0, 1)$.

Geometrically, extreme points are just the corner points of $C$.

*Example* 1.2. $\mathbb{R}^n$ is convex. Let $W$ be a subspace of $\mathbb{R}^n$ and $\mathbf{x}_1, \mathbf{x}_2 \in W$. Thus any linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$ is also in $W$, in particular the linear combination $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in W$ for $\lambda \in [0, 1]$. This shows that $W$ is convex.

*Example* 1.3. The $n$ dimensional open ball centered at $\mathbf{x}_0$ with radius $r$ is defined as

$$B_r(\mathbf{x}_0) = \{\mathbf{x} \mid |\mathbf{x} - \mathbf{x}_0| < r\}.$$

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

$$\overline{B_r(\mathbf{x}_0)} = \{\mathbf{x} \mid |\mathbf{x} - \mathbf{x}_0| \le r\}.$$

$\leq$

Both the open ball and the closed ball are convex. We prove it for the open ball. Let $\mathbf{x}_1, \mathbf{x}_2 \in B(\mathbf{x}_0)$ and $\lambda \in [0, 1]$. Then

$$\begin{aligned} |\underbrace{(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)}_{\text{convex combination}} - \mathbf{x}_0| &= |\lambda(\mathbf{x}_1 - \mathbf{x}_0) + (1-\lambda)(\mathbf{x}_2 - \mathbf{x}_0)| \\ &\le \lambda|\mathbf{x}_1 - \mathbf{x}_0| + (1-\lambda)|\mathbf{x}_2 - \mathbf{x}_0| \\ &\le \lambda r + (1-\lambda)r = r. \end{aligned}$$

$\neq$

Let $S$ be a subset of $\mathbb{R}^n$. 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 $\mathbb{R}^n - S$. Note that a boundary point can either be in $S$ or not in $S$. The set of all boundary points of $S$, denoted by $\partial S$, is the *boundary* of $S$. A set $S$ is *closed* if $\partial S \subseteq S$. A set $S$ is open if its complement $\mathbb{R}^n - S$ is closed. Note that a set that is not closed is *not* necessarily open; and a set that is not open is *not necessarily closed.* There are sets that are neither open nor closed. The *closure* of a set $S$ is the set $\overline{S} = S \cup \partial S$. The *interior* of a set $S$ is the set $S^\circ = S - \partial S$. A set $S$ is closed if and only if $S = \overline{S}$. A set $S$ is open if and only if $S = S^\circ$.

*Example* 1.4. $\mathbb{R}^n$ is both open and closed. The empty set $\emptyset$ is both open and closed.

*Example* 1.5. The hyperplane $P = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ is closed in $\mathbb{R}^n$. In fact we will show that $P \subseteq \partial P$. Without loss of generality we may assume $|\mathbf{c}| = 1$. Let $\mathbf{x} \in P$ and $B_r(\mathbf{x})$ is an open ball centered at $\mathbf{x}$ with radius $r$. Since $\mathbf{x} \in B_r(\mathbf{x})$ it remains to show that $B_r(\mathbf{x})$ contains a point not in $P$. Let

$$\mathbf{y} = \mathbf{x} + \frac{r}{2}\mathbf{c}$$

need to prove also $\partial P \subseteq P$ ($\Leftrightarrow$ if $x \notin P$ $x \notin \partial P$

then

$$\mathbf{c}^T\mathbf{y} = \mathbf{c}^T\mathbf{x} + \frac{r}{2}\mathbf{c}^t\mathbf{c} = z + \frac{r}{2} > z.$$

Hence $\mathbf{y} \notin P$. But $|\mathbf{y} - \mathbf{x}| = \frac{r}{2}$ therefore $\mathbf{y} \in B_r(\mathbf{x})$

*Example* 1.6. The half spaces

$$X_1 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \le z\} \quad \text{and} \quad X_2 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \ge z\}$$

are closed in $\mathbb{R}^n$. In fact we have $\partial X_1 = \partial X_2 =$ the hyperplane $P = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$. We will show that $\partial X_1 = P$, the proof for $\partial X_2 = P$ is similar.

( $P \subset \partial X_1$ ) Let $\mathbf{x} \in P$. For any $r > 0$, let

$$\mathbf{y}_1 = \mathbf{x} + \frac{r}{2|\mathbf{c}|}\mathbf{c}, \quad \mathbf{y}_2 = \mathbf{x} - \frac{r}{2|\mathbf{c}|}\mathbf{c}.$$

We see that $|\mathbf{x} - \mathbf{y}_1| = \frac{r}{2} = |\mathbf{x} - \mathbf{y}_2|$ so both $\mathbf{y}_1, \mathbf{y}_2 \in B_r(\mathbf{x})$. Moreover

$$\mathbf{c}^T\mathbf{y}_1 = \mathbf{c}^T\mathbf{x} + \frac{r}{2|\mathbf{c}|}\mathbf{c}^T\mathbf{c} = r + \frac{r}{2} > r$$

and therefore $\mathbf{y}_1 \notin X_1$. On the other hand

$$\mathbf{c}^T\mathbf{y}_2 = \mathbf{c}^T\mathbf{x} - \frac{r}{2|\mathbf{c}|}\mathbf{c}^T\mathbf{c} = r - \frac{r}{2} < r$$

so $\mathbf{y}_2 \in X_1$. This shows $\mathbf{x} \in \partial X_1$.

( $\partial X_1 \subset P$ ) Suppose $\mathbf{x} \notin P$. If $\mathbf{c}^T\mathbf{x} = z_1 < z$ then let $r = \frac{z-z_1}{2} > 0$. The open ball $B_r(\mathbf{x})$ lies entirely in $X_1$. So $B_r(\mathbf{x})$ contains no point outside of $X_1$, hence $\mathbf{x} \notin \partial X_1$. If $\mathbf{c}^T\mathbf{x} = z_1 > z$ then let $r = \frac{z_1-z}{2} > 0$. The open ball $B_r(\mathbf{x})$ lies entirely outside of $X_1$. So $B_r(\mathbf{x})$ contains no point of $X_1$, hence $\mathbf{x} \notin \partial X_1$. In either case $\mathbf{c} \notin \partial X_1$.

**Lemma 1.1.** *(a) All hyperplanes are convex.*

*(b) The closed half-space* $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \le z\}$ *and* $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} \ge z\}$ *are convex.*

*(c) The open half-space* $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} < z\}$ *and* $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} > z\}$ *are convex.*

*(d) Any intersection of convex sets is still convex.*

*(e) The set of all* feasible solutions *to a linear programming problem is a convex set.*

*Proof.* (a) Let $X = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ be our hyperplane. For all $\mathbf{x}_1, \mathbf{x}_2 \in X$ and $\lambda \in [0,1]$, we have

$$\mathbf{c}^T[\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2] = \lambda\mathbf{c}^T\mathbf{x}_1 + (1-\lambda)\mathbf{c}^T\mathbf{x}_2 = \lambda z + (1-\lambda)z = z.$$

Thus $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in X$. Hence $X$ is convex.

(b) and (c) can be proved similarly by replacing the equality signs in (a) by the corresponding inequality signs.

(d) Let $C = \bigcap_{\alpha\in I} C_\alpha$, where $C_\alpha$ are convex for all $\alpha$ in the index set $I$. Then for all $\mathbf{x}_1, \mathbf{x}_2 \in C$, we have $\mathbf{x}_1, \mathbf{x}_2 \in C_\alpha$ for all $\alpha \in I$. Hence for all $\lambda \in [0,1]$,

$$\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C_\alpha$$

for all $\alpha \in I$. Thus $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C$, and $C$ is convex.

(e) For any LP problem, the constraints can be written as $\mathbf{a}_i\mathbf{x} \le b_i$ or $\mathbf{a}_i\mathbf{x} = b_i$ etc. The set of points that satisfy any one of these constraints is thus a half space or a hyperplane. By (a), (b) and (c), they are convex. By (d), the intersection of all these sets, which is defined to be the set of feasible solutions, is a convex set. □

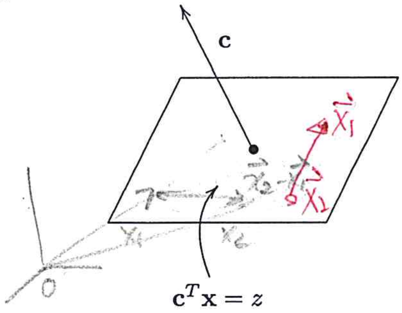
**Definition 1.7.** Let $\{x_1, \cdots x_k\}$ be a set of given points. Let

$$\mathbf{x} = \sum_{i=1}^{k} M_i\mathbf{x}_i$$

where $M_i \ge 0$ for all $i$ and $\sum_{i=1}^k M_i = 1$. Then $\mathbf{x}$ is called a *convex combination* of the points $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$.

*Example* 1.7. Consider the triangle on the plane with vertices $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$. Then any point $\mathbf{x}$ in the triangle is a convex combinations of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$. In fact, let $\mathbf{y}$ be the extension of the line segment

In $\mathbb{R}^2$, a hyperplane is given by the equation $c_1x_1 + c_2x_2 = z$, which is a straight line. In $\mathbb{R}^3$, hyperplanes are just planes.




$\vec{c}^T\vec{x} = z$

$c^T(\vec{x}_1 - \vec{x}_2) = \cancel{z} - z = 0$

Given a hyperplane $\mathbf{c}^T\mathbf{x} = z$, it is clear that it passes through the origin if and only if $z = 0$. In that case, $\mathbf{c}$ is orthogonal to every vector $\mathbf{x}$ of the hyperplane, and the hyperplane forms a vector subspace of $\mathbb{R}^n$ with dimension $n-1$.

If $z \neq 0$, then for any two vectors $\mathbf{x}_1 \neq \mathbf{x}_2$ in the hyperplane, we have

$$\mathbf{c}^T(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{c}^T\mathbf{x}_1 - \mathbf{c}^T\mathbf{x}_2 = z - z = 0.$$

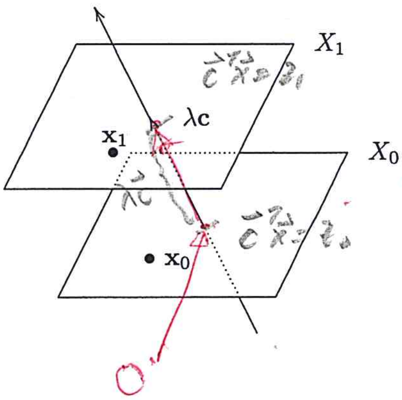
Thus $\mathbf{c}$ is orthogonal to $\mathbf{x}_1 - \mathbf{x}_2$ which is a vector lying on the hyperplane, i.e. $\mathbf{c}$ is also perpendicular to the hyperplane. We note that in this case, the hyperplane is an affine space.

**Definition 1.4.** Given the hyperplane $\mathbf{c}^T\mathbf{x} = z$, the vector $\mathbf{c}$ is called the *normal* of the hyperplane. Two hyperplanes are said to be *parallel* if their normals are parallel vectors.

*Example* 1.1. Let $\mathbf{x}_0$ be arbitrarily chosen from a hyperplane $X_0 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z_0\}$. Let $\lambda > 0$ be fixed. Then the point $\mathbf{x}_1 = \mathbf{x}_0 + \lambda\mathbf{c}$ satisfies

$$\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_0 + \lambda|\mathbf{c}|^2 = z_0 + \lambda|\mathbf{c}|^2 > z_0.$$

Here $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$. Let $z_1 \equiv z_0 + \lambda|\mathbf{c}|^2$ and define the hyperplane $X_1 = \{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z_1\}$. We see that the hyperplanes $X_0$ and $X_1$ are parallel and $X_1$ is lying in the direction of $\mathbf{c}$ from $X_0$. The distance between the hyperplanes is $\lambda$.




$z_1 = z_0 + \lambda|c|^2$

$\vec{c}^T x = z_1$

$\vec{x}_0 \in X_0$

$\vec{c}^T\vec{x} = z_0$

$\vec{x}_1 = \vec{x}_0 + \lambda\vec{c} \in X_1$

$\vec{c}^T\vec{x}_1 = \vec{c}^T\vec{x}_0 + \lambda\vec{c}^T\vec{c}$

$= z_0 + \lambda|\vec{c}|^2$

$= z_1 > z_0$

凸

## 1.7 Convex Sets

**Definition 1.5.** A set $C$ is said to be $\boxed{\textit{convex}}$ if for all $\mathbf{x}_1$ and $\mathbf{x}_2$ in $C$ and $\lambda \in [0,1]$, we have $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C$.

Geometrically, that means that given any two points in $C$, then all points on the line segment joining the given two points should also be in $C$.

$C$ $\forall\ \vec{x}_1, \vec{x}_2 \in C$ if convex combination

$\alpha\vec{x}_1 + (1-\alpha)\vec{x}_2 \in C$

$\forall\ \alpha \in [0,1]$

are all in $C$ then $C$ is convex.

Let us suppose that $c_r \neq 0$. Then

$$\sum_{j=1}^{m} \tilde{x}_{B_j} \mathbf{a}_{i_j} = \sum_{j=1}^{m} (x_{B_j} - \frac{x_{B_r}}{c_r} c_j) \mathbf{a}_{i_j} = \mathbf{b} \; .$$

Since $\tilde{x}_{B_r} = 0$, we have found a degenerate solution to $A\mathbf{x} = \mathbf{b}$, a contradiction. Thus $\{\mathbf{a}_{i_j}\}_{j=1}^{m}$ is linearly independent. Using the fact that $x_{B_j} \neq 0$ for all $j$, we can, by replacement method, show that any $m-1$ vectors of $\{\mathbf{a}_{i_j}\}_{j=1}^{m}$ together with the vector $\mathbf{b}$ are also linearly independent. In fact, for all $x_{B_r} \neq 0$, we have

$$-\sum_{\substack{j=1 \\ j \neq r}}^{m} \frac{x_{B_j}}{x_{B_r}} \mathbf{a}_{i_j} + \frac{1}{x_{B_r}} \mathbf{b} = \mathbf{a}_{i_r} .$$

+

Using similar argument as above, we conclude that the set

$$\{\mathbf{a}_{i_1}, \cdots \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots \mathbf{a}_{i_m}, \mathbf{b}\}$$

is also linearly independent.
($\Leftarrow$) Let $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}\}$ be an arbitrary set of $m$ columns of $A$. By assumption, it is linearly independent. Since $\mathbf{a}_{i_j}$ are $m$-vectors, we see that the set $\{\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}, \mathbf{b}\}$ is linearly dependent. Hence there exist $x_{B_j}$, $j = 1, \cdots, m$ such that

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

Thus basic solution exists for such choice of $m$ columns of $A$. Next we claim that $x_{B_j} \neq 0$ for all $j = 1, 2, \cdots, m$. Suppose that $x_{B_r} = 0$. Then

$$\mathbf{b} - \sum_{\substack{j=1 \\ j \neq r}}^{m} x_{B_j} \mathbf{a}_{i_j} = \mathbf{0} \; .$$

That means that the set

$$\{\mathbf{a}_{i_1}, \cdots \mathbf{a}_{i_{r-1}}, \mathbf{a}_{i_{r+1}}, \cdots \mathbf{a}_{i_m}, \mathbf{b}\}$$

is linearly dependent, hence a contradiction to our assumption. Thus $x_{B_j} \neq 0$ for all $j = 1, \cdots, m$.
□

By the same arguments used in the proof of the above theorem, we have the following corollary.

**Corollary 1.1.** *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 linearly independence of $\mathbf{b}$ with every $m-1$ columns of $\{\mathbf{a}_{i_1}, \cdots, \mathbf{a}_{i_m}\}$.*

## 1.6 Hyperplanes

**Definition 1.3.** A *hyperplane* in $\mathbb{R}^n$ is defined to be the set of all points in $\{\mathbf{x} \mid \mathbf{c}^T\mathbf{x} = z\}$ where $\mathbf{c}$ is a fixed nonzero vector in $\mathbb{R}^n$ and $z \in \mathbb{R}$.

$c_1x_1 + c_2x_2 + \cdots + c_nx_n = z$

Thus $\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ is also a solution. Hence we have proved that if $A\mathbf{x} = \mathbf{b}$ has more than one solution, then it has infinite many solutions.

To characterize these solutions, let us first suppose that $A$ is an $m$-by-$n$ matrix with $\text{rank}(A) = m < n$. Then $A\mathbf{x} = \mathbf{b}$ can be written as

$$B\mathbf{x}_B + R\mathbf{x}_\beta = \mathbf{b} \ ,$$

$A\vec{x} = \vec{b}$

$m[\underset{m}{B} \mid \underset{n-m}{R}]\begin{bmatrix}\vec{X}_B \\ \vec{X}_\beta\end{bmatrix} = \vec{b}$

linearly independent

where

$$B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_m],$$

$$R = \begin{bmatrix} a_{1m+1} & a_{1m+2} & \cdots & a_{1n} \\ a_{2m+1} & a_{2m+2} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{mn+1} & a_{mn+2} & \cdots & a_{mn} \end{bmatrix} = [\mathbf{a}_{m+1}, \mathbf{a}_{m+2}, \cdots, \mathbf{a}_n],$$

$$\mathbf{x}_B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \text{and} \quad \mathbf{x}_\beta = \begin{bmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{bmatrix} .$$

$B\vec{X}_B = \vec{b} - \underbrace{R\vec{X}_\beta}$

Since the ordering of the variables $x_i$ are irrelevant, we can assume that the variables have been reordered so that $B$ is nonsingular. Or equivalently, we can consider the nonsingular $B$ as being formed by suitably picking $m$ columns of $A$. Then we have

$$\boxed{\mathbf{x}_B = B^{-1}(\mathbf{b} - R\mathbf{x}_\beta) \ .} \tag{1.2}$$

Hence given any $\mathbf{x}_\beta$ we can solve uniquely for $\mathbf{x}_B$ in terms of $\mathbf{x}_\beta$. Thus to find the solution to (1.1), we can assign arbitrary values to the $(n-m)$ variables in $\mathbf{x}_\beta$ and determine from (1.2) the values for the remaining $m$ variables in $\mathbf{x}_B$. Then $\mathbf{x} = \begin{bmatrix}\mathbf{x}_B \\ \mathbf{x}_\beta\end{bmatrix}$ is a solution to $A\mathbf{x} = \mathbf{b}$.

$\vec{b} = \vec{0}$

## 1.4 Homogeneous Systems of Linear Equations

A homogeneous system of linear equations is one that is of the form $A\mathbf{x} = \mathbf{0}$. Its solutions form an $(n-m)$ dimensional subspace of $\mathbb{R}^n$. In fact, the solution space is just the kernel of the linear mapping represented by the matrix $A$. Since $A$ is a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, we see that

$X_\beta \in \mathbb{R}^{n-m}$

$B\vec{X}_B = -R\vec{X}_\beta \Leftrightarrow A\vec{x} = \vec{0}$

$$\begin{aligned} \text{dimension of kernel of } A &= \text{dimension of } \mathbb{R}^n - \text{dimension of range of } A \\ &= n - \text{rank}(A) = n - m. \end{aligned}$$

$A\vec{x} = \vec{0}$

$[B \mid R]\begin{bmatrix}\vec{X}_B \\ \vec{X}_R\end{bmatrix} = \vec{0}$

Note that if $\mathbf{x}_1$ is a solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{x}_0 \neq \mathbf{0}$ is a solution of $A\mathbf{x} = \mathbf{0}$, then $\mathbf{x}_1 + \mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{b}$. In fact,

$$A(\mathbf{x}_1 + \mathbf{x}_0) = A\mathbf{x}_1 + A\mathbf{x}_0 = \mathbf{b} + \mathbf{0} = \mathbf{b} \ .$$

$\vec{X}_B = B^{-1}(-R\vec{X}_\beta)$

Clearly $\mathbf{x}_1 \neq \mathbf{x}_1 + \mathbf{x}_0$. Hence we have found two distinct solutions to $A\mathbf{x} = \mathbf{b}$ and by the results in §3, we see that $A\mathbf{x} = \mathbf{b}$ has infinite many solutions. Thus we have proved that if $n > m$, and if $A\mathbf{x} = \mathbf{b}$ is not inconsistent, then $A\mathbf{x} = \mathbf{b}$ has infinite many solutions.

$\vec{X}_\beta \in \mathbb{R}^{n-m}$

We remark that if $\mathbf{b} \neq \mathbf{0}$, the solutions of $A\mathbf{x} = \mathbf{b}$ do not form a subspace of $\mathbb{R}^n$, but is a space translated away from the origin. Such a space is called an *affine space.*

Kernel of $A = \{\vec{x} \mid A\vec{x} = \vec{0}\}$ $\quad \dim(\mathrm{Ker}(A)) = n-m$

## 1.5 Basic Solutions = Corner Points $A\vec{x}=\vec{b} \Leftrightarrow [B\,|\,R]\begin{bmatrix}\vec{x}_B\\ \vec{x}_\beta\end{bmatrix}=\vec{b}$

**Definition 1.1.** Consider $A\mathbf{x} = \mathbf{b}$ with rank$(A) = m < n$. Let $B$ be a matrix formed by choosing $m$ columns out of the $n$ columns of $A$. If $B$ is a non-singular matrix and if all the $(n-m)$ variables not associated with these columns are set equal to zero, then the solution to the resulting system of equations is called a *basic solution*. We call the $m$ variables $x_i$ associated with these $m$ columns the *basic variables*, the other variables are called the *non-basic variables*. $\vec{x}_\beta=\vec{0}$

$\vec{x}_\beta=\vec{0}$

$B\vec{x}_B=\vec{b}$

$\vec{x}_B=B^{-1}\vec{b}$

Using (1.2), we see that $\mathbf{x}_B$ are the basic variables and $\mathbf{x}_\beta$ are the non-basic variables. To get a basic solution, we set $\mathbf{x}_\beta = \mathbf{0}$. Then $\mathbf{x}_B = B^{-1}\mathbf{b}$ and $\mathbf{x} = \begin{bmatrix} B^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$ is a basic solution.

**Definition 1.2.** A basic solution to $A\mathbf{x} = \mathbf{b}$ is *degenerate* if one or more of the $m$ basic variables vanish.

*Example* 1.1. Consider the system $A\mathbf{x} = \mathbf{b}$ where

$$\begin{pmatrix}1&2\\0&1\end{pmatrix}(\vec{x}_B)=\begin{pmatrix}1\\1\end{pmatrix} \qquad A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \qquad \begin{pmatrix}-1\\1\end{pmatrix}=\begin{pmatrix}1&2\\0&1\end{pmatrix}^{-1}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1&-2\\0&1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}$$

If we take $B = [\mathbf{a}1\ \mathbf{a}_2] = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. In this case $\mathbf{x}_B = [-1, 1]^T$ and $\mathbf{x} = [-1, 1, 0, 0]^T$ is the corresponding basic solution.

Basic non basic non basic

We can also take $B = [\mathbf{a}_1\ \mathbf{a}_4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $\mathbf{x}_B = [1, 1]^T$ and $\mathbf{x} = [1, 0, 0, 1]^T$ is the corresponding basic solution.

$\begin{pmatrix}1&0\\0&1\end{pmatrix}\vec{x}_B=\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$ Basic variable

For $B = [\mathbf{a}_2\ \mathbf{a}_3] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ we have $\mathbf{x}_B = [1/3, 1/3]^T$. Hence $\mathbf{x} = [0, 1/3, 1/3, 0]^T$ is the corresponding basic solution.

2nd 3rd $\begin{pmatrix}1&2\\2&1\end{pmatrix}\vec{x}_B=\begin{pmatrix}1\\1\end{pmatrix}$

Note that all these basic solutions are non degenerate.

Thus if any one of the elements of $\mathbf{x}_B$ is zero, the basic solution is degenerate.

Although the number of solutions to $A\mathbf{x} = \mathbf{b}$ are in general infinite, we will see later that the optimal solutions in LP problems are basic solutions. Therefore, we will like to know how many basic solutions are there. This is equivalent to asking how many such nonsingular matrices $B$ can possibly be formed from the columns of $A$. It is obvious that the number of such matrices is bounded by $C_m^n = \dfrac{n!}{m!(n-m)!}$.

$\begin{pmatrix}1&1\\1&1\end{pmatrix}$

A b

$\begin{bmatrix}1&1&1&0\\0&1&2&1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$

$A\vec{x}=\vec{b}$

$\vec{x}_\beta=(x_3\ x_4)$

$\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$

$\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$

$\vec{x}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix}$ non basic

**Theorem 1.1.** *A necessary and sufficient condition for the existence and non-degeneracy of all possible basic solutions of $A\mathbf{x} = \mathbf{b}$ is the linearly independence of every set of $m$ columns of the augmented matrix $A_b = [A, \mathbf{b}]$.*

*Proof.* ($\Rightarrow$) Suppose that all basic solutions exist and are not degenerate. Then for any set of $m$ columns of $A$, say

$$\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \cdots, \mathbf{a}_{i_m}, \qquad 1 \le i_j \le n, 1 \le j \le m,$$

there exists $x_{B_j}$, $j = 1, 2, \cdots, m$ such that

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

Since the solution is non degenerate, all $x_{B_j} \neq 0$. We first claim that $\{\mathbf{a}_{i_j}\}_{j=1}^m$ is linearly independent. For if not, then there exist $c_j$, $j = 1, \cdots, m$ not all zeros, such that

$$\sum_{j=1}^{m} c_j \mathbf{a}_{i_j} = \mathbf{0}.$$