Farkas' Lemma

This Hypercard stack was prepared by:
Dennis L. Bricker,
Dept. of Industrial Engineering,
University of Iowa,
Iowa City, Iowa 52242
e-mail: dbricker@icaen.uiowa.edu

author

Farkas' Lemma

Let
$A \in R^{m\times n}$, *i.e., A is mxn matrix,*
$b \in R^m$,
$x \in R^n$, $y \in R^m$

The following statements are equivalent:

1 $y^T A \le 0 \Rightarrow y^T b \le 0$

&

2 $\exists\ x$ such that $A x = b,\ x \ge 0$

Proof

Interpretation

Application

Proof

Consider the primal/dual LP pair:

P Minimize $0x$
subject to $A\,x = b$
$x \ge 0$

D Maximize $y^T b$
subject to $A^T y \le 0$,
i.e., $y^T A \le 0$

Problem **D** is feasible (e.g., let y=0, for which the objective $y^T b$ is zero.)

If statement **1** is true, i.e., $y^T A \le 0 \Rightarrow y^T b \le 0$ then y=0 must be optimal for problem **D**.

If y=0 is optimal for **D**, then by LP duality theory, **P** is feasible (with optimal value 0), proving that **1** ⇒ **2**.

Suppose that $Ax=b$ for some $x\ge 0$, and $y^T A \le 0$ for some y.

Then $y^T A \le 0 \Rightarrow y^T A x \le 0 \Rightarrow y^T b \le 0$

proving that **2** ⇒ **1**.

QED

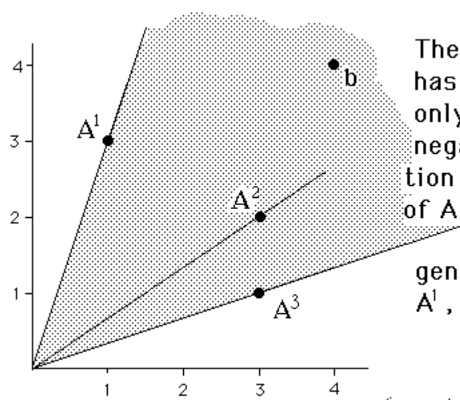
GEOMETRIC ILLUSTRATION OF FARKAS' LEMMA

Let $A = \begin{bmatrix} 1 & 3 & 3 \\ 3 & 2 & 1 \end{bmatrix}$, $b = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

The columns of A are points (vectors) in R^2

$$A^1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix},\ A^2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix},A^1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix},\ b = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

(requirements space)



The system Ax=b has a solution if & only if b is a non-negative combination of the columns of A, i.e., iff b lies in the *cone* generated by A^1, A^2, & A^3

(requirements space)

For example,

$$A^1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix},\ A^2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix},A^1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

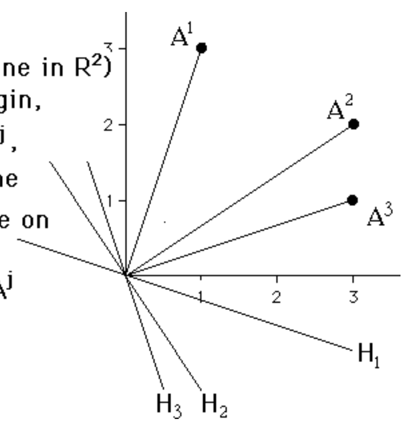
$$b = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 1\ A^1 + 0\ A^2 + 1\ A^3$$

$$= \frac{4}{7} A^1 + \frac{4}{7} A^2 + 0\ A^3$$

$$= \frac{11}{14} A^1 + \frac{4}{7} A^2 + \frac{1}{2} A^3$$

..., etc.

Let H_j be the hyperplane (a line in R^2) through the origin, orthogonal to A^j, and let H_j^- be the closed halfspace on the side of H_j not containing A^j



$y^T A^j = 0 \Leftrightarrow y \perp A^j$

$\Leftrightarrow y \in H_j$

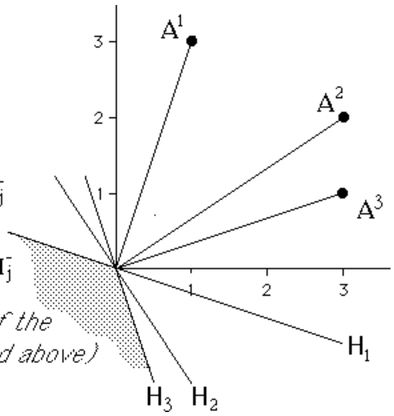
Also,

$y^T A^j \le 0 \Leftrightarrow y \in H_j^-$

Therefore,

$y^T A \le 0 \Leftrightarrow y \in \bigcap_j H_j^-$

(the intersection of the half-spaces, shaded above)



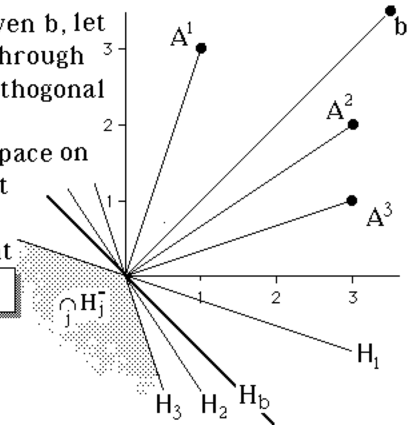
Likewise, for a given b, let H_b = hyperplane through the origin, orthogonal to b, and H_b^- = closed halfspace on side of H_b not containing b.

Then the statement

$y^T A \le 0 \Rightarrow y^T b \le 0$

simply says that

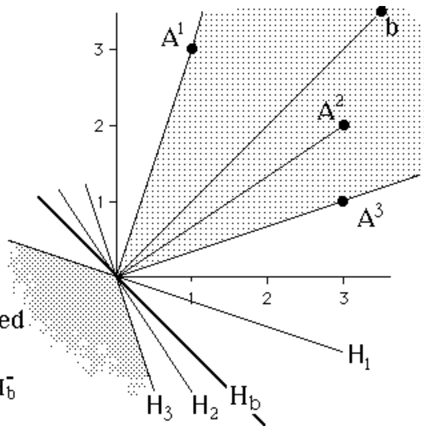
$\bigcap_j H_j^- \subseteq H_b^-$



EXAMPLE 1

$b = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

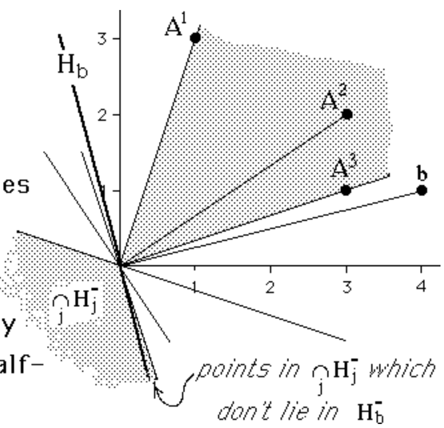
Note that b is in the cone generated by $A^1, A^2, \& A^3$ and that $\bigcap_j H_j^- \subseteq H_b^-$



EXAMPLE 2

$b = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

In this case, the vector b does not lie in the cone generated by A, nor does $\bigcap_j H_j^-$ lie entirely in the closed half-space H_b^-



APPLICATION TO NONLINEAR PROGRAMMING

Consider the problem

Minimize $f(x)$
subject to $g_i(x) \le 0$, $i=1,2,\ldots m$

Denote

$b \equiv -\nabla f(x^*)$

$A^i \equiv \nabla g_i(x^*)$

$y \equiv d$ *(direction vector)*

$x_i \equiv \lambda_i$ for $i \in I \equiv \{ i \mid g_i(x^*) = 0 \}$

(Lagrange multiplier) *(index set of tight constraints)*

Farkas' Lemma

1 $y^T A \le 0 \Rightarrow y^T b \le 0$

&

2 $\exists\, x$ such that $A x = b$, $x \ge 0$

are equivalent statements

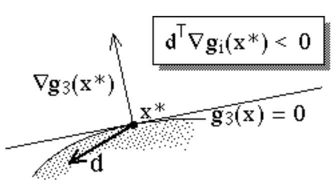
That is,

1 $d^T \nabla g_i(x^*) \le 0 \;\forall i \in I \Rightarrow -d^T \nabla f(x^*) \le 0$

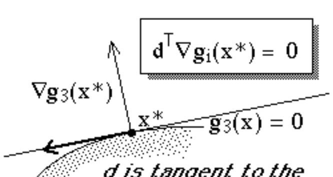
&

2 $\exists\, \lambda_i \ge 0$ such that $\sum_{i \in I} \lambda_i \nabla g_i(x^*) = -\nabla f(x^*)$

are equivalent statements

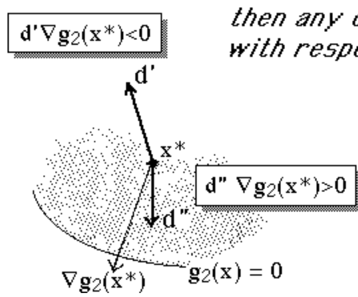


d is a feasible direction



d is tangent to the constraint boundary

d is "feasible", but any positive step in this direction may be infeasible



If a constraint is not tight, then any direction is feasible with respect to that constraint!

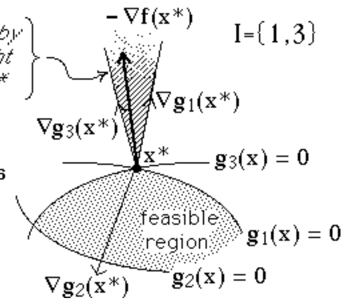
1 $d^T \nabla g_i(x^*) \le 0 \ \forall i \in I \Rightarrow -d^T \nabla f(x^*) \le 0$

directions satisfying $d^T \nabla g_i(x^*) \le 0 \ \forall i \in I$
are feasible directions

directions satisfying $d^T \nabla f(x^*) \ge 0$
are directions of ascent

1 Every feasible direction is non-improving

2 $\exists\ \lambda_i \ge 0$ such that $\sum_{i \in I} \lambda_i \nabla g_i(x^*) = -\nabla f(x^*)$



Steepest descent direction
is in the cone generated by
gradients of tight constraints

Karush-Kuhn-Tucker
Conditions

K-K-T "Necessary" Condition for Optimality

If x^* is an optimal solution to

Minimize $f(x)$
subject to $g_i(x) \le 0$, $i=1,2,\ldots m$

then

The directional derivative of $f(x)$ is nonnegative in every feasible direction at x^*

K-K-T "Necessary" Condition for Optimality

If x^* is an optimal solution to

Minimize $f(x)$
subject to $g_i(x) \le 0$, $i=1,2,\ldots m$

then

The steepest descent direction at x^* is in the cone generated by the gradients of the tight constraints at x^*

Equivalent condition, according to Farkas' lemma