Week 5

Note Title 4/10/2021

Recall: if we are interested in solving a linear system $Ax=b$ then the linear map $x \mapsto Ax$, represented by the tableau

$$\begin{array}{cc} & x \\ y = & \boxed{\;A\;} \end{array}$$

is relevant.

Of course, we may use the array of numbers in A to define many other maps, e.g.

$$\begin{array}{cc} & u \\ v = & \boxed{\;A^T\;} \end{array}, \quad \begin{array}{cc} & u \\ v = & \boxed{\;-A^T\;} \end{array}, \quad \text{or maybe} \quad \begin{array}{cc} & x \\ y = & \boxed{\;A.\hat{}2\;} \end{array}$$

square every entry of A

but there is no reason to believe that these maps would have any meaningful relationship with the original $y=Ax$ map.

As it turns out, there is a miraculous relationship between A and A^T.

The way we begin to describe this relationship is through the following fact:

Theorem 4.1.1 (Dual Transformation). *A Jordan exchange with pivot element A_{rs} has two equivalent interpretations:*

1. *(Primal): Solve $y_r = \sum_{j=1}^{n} A_{rj}x_j$ for x_s and substitute for x_s in the remaining $y_i = \sum_{j=1}^{n} A_{ij}x_j$, $i \neq r$.*
2. *(Dual): Solve $v_s = -\sum_{i=1}^{m} A_{is}u_i$ for u_r and substitute for u_r in the remaining $v_j = -\sum_{i=1}^{m} A_{ij}u_i$, $j \neq s$.*

In other words :

[see computer demo]

$$\begin{array}{cc} & x \\ y = & \boxed{A} \end{array} \qquad\qquad \begin{array}{cc} & u \\ v = & \boxed{-A^T} \end{array}$$

$$\Big\downarrow \text{jx}(r,s) \qquad\qquad\qquad \Big\downarrow \text{jx}(s,r)$$

$$\begin{array}{cc} & x_1 \cdots y_r \cdots x_n \\ \begin{array}{c} y_1 = \\ \vdots \\ x_s = \\ \vdots \\ y_m = \end{array} & \boxed{\quad B \quad} \end{array} \qquad\qquad \begin{array}{cc} & u_1 \cdots v_s \cdots u_m \\ \begin{array}{c} v_1 = \\ \vdots \\ u_r = \\ \vdots \\ v_n = \end{array} & \boxed{\quad ? \quad} \end{array} \neq -B^T \;!!$$

Proof: WLOG, we prove the theorem for $r=m$, $s=n$, i.e. assuming $\alpha = A_{mn} \neq 0$ is the pivot.

Write

$$A = \begin{bmatrix} \tilde{A} & a \\ \rho & \alpha \end{bmatrix}$$

write $x = \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} \tilde{x} \\ x_n \end{bmatrix}$ $\quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} = \begin{bmatrix} \tilde{y} \\ y_m \end{bmatrix}$

Recall what "$jx(m,n)$" means:

$$\begin{cases} \tilde{y} = \tilde{A}\tilde{x} + a\, x_n \\ y_m = \rho\tilde{x} + \alpha\, x_n \end{cases} \Rightarrow \begin{cases} \tilde{y} - a x_n = \tilde{A}\tilde{x} \\ -\alpha x_n = \rho\tilde{x} - y_m \end{cases}$$

$$\Updownarrow$$

$$\begin{bmatrix} I & -a \\ 0 & -\alpha \end{bmatrix}\begin{bmatrix} \tilde{y} \\ x_n \end{bmatrix} = \begin{bmatrix} \tilde{A} & 0 \\ \rho & -1 \end{bmatrix}\begin{bmatrix} \tilde{x} \\ y_m \end{bmatrix}$$

so $B = \begin{bmatrix} I & -a \\ 0 & -\alpha \end{bmatrix}^{-1}\begin{bmatrix} \tilde{A} & 0 \\ \rho & -1 \end{bmatrix} = \begin{bmatrix} I & -\alpha^{-1}a \\ 0 & \alpha^{-1} \end{bmatrix}\begin{bmatrix} \tilde{A} & 0 \\ \rho & -1 \end{bmatrix} = \begin{bmatrix} \tilde{A} - \alpha^{-1}a\rho & \alpha^{-1}a \\ -\alpha^{-1}\rho & \alpha^{-1} \end{bmatrix}$

$$\iff [v_1, \dots, v_n] = -[u_1, \dots, u_m] A$$

Now, for $jx(n,m)$ on $V = -A^T u$:

write $$[v_1, \dots, v_{n-1}, v_n] = [\tilde{v}, v_n]$$

$$[u_1, \dots, u_{m-1}, u_m] = [\tilde{u}, u_m]$$

$$[\tilde{v}, v_n] = -[\tilde{u}, u_m] \begin{bmatrix} \hat{A} & a \\ r & \alpha \end{bmatrix}$$

$$\begin{cases} \tilde{v} = -\tilde{u}\hat{A} - u_m r \\ v_n = -\tilde{u} a - u_m \alpha \end{cases} \Rightarrow \begin{cases} \tilde{v} + u_m r = -\tilde{u}\hat{A} \\ u_m \alpha = -\tilde{u} a - v_n \end{cases}$$

or

$$[\tilde{v} \;\; u_m] \begin{bmatrix} I & 0 \\ r & \alpha \end{bmatrix} = -[\tilde{u} \;\; v_n] \begin{bmatrix} \hat{A} & a \\ 0 & 1 \end{bmatrix}$$

so $$[\tilde{v}, u_m] = -[\tilde{u}, v_n] \underbrace{\begin{bmatrix} \hat{A} & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ r & \alpha \end{bmatrix}^{-1}}$$

$$= \begin{bmatrix} \hat{A} & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ -\alpha^{-1} r & \alpha^{-1} \end{bmatrix} = \begin{bmatrix} \hat{A} - \alpha^{-1} a r & a\alpha^{-1} \\ -\alpha^{-1} r & \alpha^{-1} \end{bmatrix} = B \quad \blacksquare$$

A clever notation:

use

$$\begin{array}{cc} & v= \\ & x \\ -u,\ y= & \boxed{A} \end{array}$$

to simultaneously represent the two maps/tableaus

$$\begin{array}{cc} & x \\ y= & \boxed{A} \end{array} \quad \text{and} \quad \begin{array}{cc} & u \\ v= & \boxed{-A^T} \end{array} = \begin{array}{cc} & v= \\ -u & \boxed{A} \end{array}$$

$v=-A^Tu$

$\Leftrightarrow v^T=(-u^T)A$

Then, by theorem 4.1.1, any (valid) jordan exchange on the "doubly decorated" tableau

$$\begin{array}{cc} & v= \\ & x \\ -u,\ y= & \boxed{A} \end{array} \xrightarrow{jx(r,s)} \begin{array}{cc} & \tilde{v}= \\ & \tilde{x} \\ -\tilde{u},\ \tilde{y}= & \boxed{B} \end{array}$$

correctly maintain the linear relationships between the x-y variables and the u-v variables.

See Example 4-1-1.

This "doubly decorated tableau" idea (based on Theorem 4.1.1) gives an interesting proof of the "row rank = column rank" theorem in linear algebra. Here is the idea:

Recall: **Theorem 4.1.3.** *Given $A \in \mathbf{R}^{m\times n}$, form the tableau $y := Ax$. Using Jordan exchanges, pivot as many of the y's to the top of the tableau as possible. The rank of A is equal to the number of y's pivoted to the top.*

$$
\begin{array}{r|c|}
 & v = \\
 & x \\
\hline
-u,\ y = & A \\
\hline
\end{array}
\quad \xrightarrow[\substack{\text{the top, and} \\ \text{reorder the rows} \\ \text{and columns}}]{\substack{\text{exchange as} \\ \text{many } y_i\text{'s to}}} \quad
\begin{array}{r|cc|}
 & u_{I_1} = & v_{J_2} = \\
 & y_{I_1} & x_{J_2} \\
\hline
-v_{J_1}\ \ x_{J_1} = & B_{I_1 J_1} & B_{I_1 J_2} \\
-u_{I_2}\ \ y_{I_2} = & B_{I_2 J_1} & 0 \\
\hline
\end{array}
$$

$|I_1|$ = max number of linearly independent rows ← row rank of A

$\|$ max number of linearly independent rows of $-A^T$

$\|$

maximum number linearly independent columns of A ← column rank of A

$\boxed{\text{Why would } A^T \text{ have anything to do with anything we care?}}$

Consider the diet problem from Week 1:

One seeks the diet with the lowest cost that achieves all the nutritional requirements:

$$\min\ p_1x_1+\cdots+p_nx_n \quad \text{s.t.} \quad \begin{aligned} A_{11}x_1+\cdots+A_{1n}x_n &\geq b_1 \\ &\vdots \\ A_{m1}x_1+\cdots+A_{mn}x_n &\geq b_m \\ x_1,\ldots,x_n &\geq 0 \end{aligned}$$

p_j = cost of 1 unit of food j

x_j = # of units of food j consumed

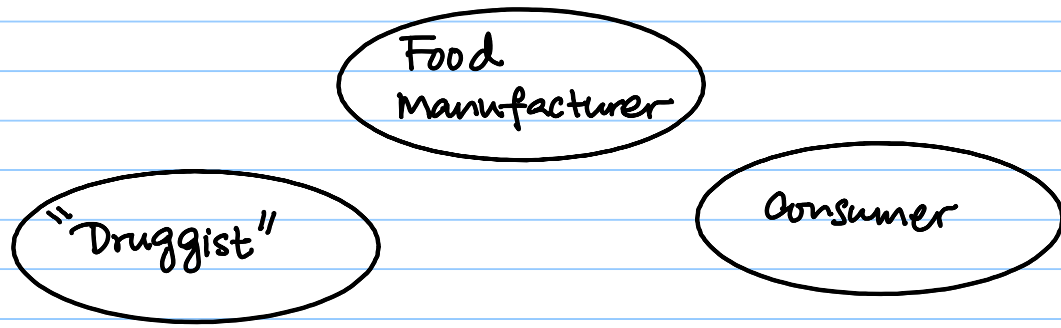
b_i = minimum requirement of nutrient i

A_{ij} = amount of nutrient i in 1 unit of food j

From the view of such a customer, the only concern is the cost of buying the required nutrients.

what if someone can directly sell you (or the food manufacturers) the nutrients?

y_i = price of 1 unit of nutrient i.



The "druggist", who sells the nutrients directly, feels that no one would buy the "raw nutrients" from him if he charges more than what one has to pay for the equivalence of nutrients from any food. This means $y_1, \dots, y_m$ should satisfy:

$$[y_1, \cdots, y_m] \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \leq [P_1, \cdots, P_n]$$

Note: $\sum_{i=1}^{m} y_i A_{ij}$ = price of the equivalence of nutrients in 1 unit of food j.

(see how A^T shows up above !)

within the "no more expensive than food" constraints, and knowing the customers need b_i units of nutrient i, $\forall i=1,\dots,m$, the druggist maximizes his revenue by solving :

$$\max_{y_1,\dots,y_m} \; b_1 y_1 + \cdots + b_m y_m \quad \text{s.t.}$$

$$[y_1, \dots, y_m] \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \leq [p_1, \dots, p_n]$$

The consumer's diet problem :

$$\boxed{\begin{aligned} &\min_{x \in \mathbb{R}^n} \; p^T x \\ &\text{s.t. } Ax \geq b \\ &\qquad\; x \geq 0 \end{aligned}}$$

The druggist's problem :

$$\boxed{\begin{aligned} &\max_{y \in \mathbb{R}^m} \; b^T y \\ &\text{s.t. } \; y^T A \leq p^T \\ &\qquad\quad y \geq 0 \end{aligned}} \iff A^T y \leq p$$

We call the LP on the right the *dual* of the LP on the left. We refer to these two problems as a *primal-dual pair* of LPs.

Note: The dual of the dual is the primal.

Here is why:

$$
\begin{tikzcd}[column sep=huge, row sep=large]
\begin{array}{ll} \min\limits_{x} & p^T x \\ \text{s.t.} & Ax \geq b \\ & x \geq 0 \end{array} \arrow[r, "\text{dual}"] & \begin{array}{ll} \max\limits_{y} & b^T y \\ \text{s.t.} & A^T y \leq p \\ & y \geq 0 \end{array} \arrow[d, "\text{standardize}"] \\
\begin{array}{ll} \max\limits_{x \in \mathbb{R}^n} & -p^T x \\ \text{s.t.} & (-A^T)^T x \leq -b \\ & x \geq 0 \end{array} \arrow[u, "\text{standardize}"] & \begin{array}{ll} \min & -b^T y \\ \text{s.t.} & -A^T y \geq -p \\ & y \geq 0 \end{array} \arrow[l, "\text{dual}"']
\end{tikzcd}
$$

Note: $-(b^T y) = (-b)^T y$

$(-A^T)^T = -A$

A primal-dual pair of LP can be represented simultaneously by a single "doubly decorated" tableau:

dual slack variables

dual objective

$$
\begin{array}{ccc|ccc|c}
 & & & u_{m+1} = & \cdots & u_{m+n} = & w = \\
 & & & x_1 & \cdots & x_n & 1 \\
\hline
-u_1 & x_{n+1} & = & A_{11} & \cdots & A_{1n} & -b_1 \\
 & \vdots & & \vdots & \ddots & \vdots & \vdots \\
-u_m & x_{n+m} & = & A_{m1} & \cdots & A_{mn} & -b_m \\
\hline
1 & z & = & p_1 & \cdots & p_n & 0 \\
\hline
\end{array}
$$

primal slack variables

primal objective

In virtue of Theorem 4.1.1, a jordan exchange of this tableau would maintain the relationships of both the primal and dual variables.

See Example 4-2-1.