I.C. Matrix algebra

Before formalizing Gauss-Jordan in terms of a fixed procedure for row-reducing A, we briefly review some properties of matrix multiplication. Let

$$m\{\ \underbrace{[A_{ij}]}_{n}\ \ ,\ \ n\{\ \underbrace{[B_{jk}]}_{p}\ \ ,\ \ p\{\ \underbrace{[C_{kl}]}_{q}$$

be matrices, with entries in (say) $\mathbb{R}$, or more generally any "field" (cf. §II.A). The *transpose* of A, written ${}^tA$, is the $n \times m$ matrix with entries

$$({}^tA)_{ij} = A_{ji},$$

and A is *symmetric* if ${}^tA = A$. (I write the superscript "t" on the left so that later we can talk about the transpose inverse ${}^tA^{-1}$ without parentheses.) The $m \times m$ "identity" matrix, denoted by $\mathbb{I}_m$ (or just $\mathbb{I}$), is defined to have $(i,j)^{\text{th}}$ entry δ_{ij} (i.e. 1 if $i = j$ and 0 otherwise). We will also write $\mathbb{E}_{ij}^m$ (or just $\mathbb{E}_{ij}$) for the matrix with $(i,j)^{\text{th}}$ entry 1 and all other entries zero.

Multiplication: At any rate, we define the matrix product AB to be the $m \times p$ matrix with entries

$$(AB)_{ik} := \sum_{j=1}^{n} A_{ij}B_{jk.} \tag{I.C.1}$$

Associativity of this product follows from associativity of the ground field:

$$((AB)C)_{il} := \sum_k (AB)_{ik}C_{kl} = \sum_k (\sum_j A_{ij}B_{jk})C_{kl} = \sum_{j,k}(A_{ij}B_{jk})C_{kl}$$
$$= \sum_{j,k} A_{ij}(B_{jk}C_{kl}) = \ldots = (A(BC))_{il}.$$

On the other hand, commutativity **fails** for matrix multiplication: that is, $AB \neq BA$ in general. Indeed, BA isn't even defined unless $p = m$, in which case the closest true statement one has in general is

$$BA = {}^t({}^tA\,{}^tB).$$

A simple example of noncommutativity for square matrices is

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

For a physicist, noncommutativity is essential, since it's the entire point of the Heisenberg uncertainty principle that the position and momentum operators don't commute! Or as the Mad Hatter says, "seeing what you eat" and "eating what you see" are not at all the same thing.[2]

Inverses: if A is $m \times n$ for $m < n$, it cannot have a left inverse (L such that $LA = \mathbb{I}_n$) but may have many right inverses (R such that $AR = \mathbb{I}_m$). An example, where a, b can be any real numbers:

$$\overbrace{\begin{bmatrix} 1 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{A} \cdot \begin{bmatrix} 1-at & -bt \\ a & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If $m > n$ then the situation is just reversed. For square matrices ($m = n$) we will prove in §I.D that existence of a left inverse is equivalent to existence of a right inverse. But if *both* exist for a matrix A, they must be the same:

$$BA = \mathbb{I},\ AC = \mathbb{I} \implies B = B\mathbb{I} = B(AC) = (BA)C = \mathbb{I}C = C.$$

In this situation we say A is *invertible*, denoting the (left and right) inverse matrix by A^{-1}. Products and inverses of invertible matrices are invertible; e.g. for products AB, $B^{-1}A^{-1}$ furnishes a 2-sided inverse.

I.C.2. EXAMPLE. If $\alpha\delta - \beta\gamma \neq 0$ (for $\alpha, \beta, \gamma, \delta$ in your favorite field), we have

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} = \frac{1}{\alpha\delta - \beta\gamma} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}.$$

[2]though putting these words in the Mad Hatter's mouth may have been a polemic on Lewis Carroll's part against the quaternions.

Vectors and matrix multiplication: For $\vec{x} \in \mathbb{R}^n$, here are some characterizations of the matrix-vector product $A\vec{x}$ in terms of rows and columns of A:

$$A\vec{x} = \begin{pmatrix} \leftarrow & \vec{r}_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \vec{r}_m & \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow \\ \vec{x} \\ \downarrow \end{pmatrix} \overset{(*)}{=} \begin{pmatrix} \vec{r}_1 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{pmatrix}$$

$$= \begin{pmatrix} \uparrow & & \uparrow \\ \vec{c}_1 & \dots & \vec{c}_n \\ \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \overset{(**)}{=} x_1\vec{c}_1 + \ldots + x_n\vec{c}_n = \sum_{j=1}^{n} x_j\vec{c}_j.$$

Here $(*)$ is directly from the definition (I.C.1) of matrix multiplication, while $(**)$ is checked in Exercise (1) below.

In particular, writing $\hat{e}_j$ for the jth coordinate vector of $\mathbb{R}^n$ (i.e. the jth column of $\mathbb{I}_n$), we see that $A\hat{e}_j = \vec{c}_j$ and so

$$A = \begin{pmatrix} \uparrow & & \uparrow \\ A\hat{e}_1 & \dots & A\hat{e}_n \\ \downarrow & & \downarrow \end{pmatrix}.$$

(With this understood, you should now be able to easily convince yourself that the columns of a matrix product AB are linear combinations of the columns of A! See Exercise (4).)

There are two different ways to multiply vectors as matrices:

$$\begin{bmatrix} 2 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} = 4 = \text{ dot (interior) product,}$$

$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & -2 \\ -3 & -1 & 1 \\ 3 & 1 & -1 \end{bmatrix} = \text{ exterior product.}$$

If $\vec{x}$ and $\vec{y}$ are two column vectors, then the dot product $\vec{x} \cdot \vec{y}$ in terms of matrix multiplication is ${}^tX \cdot Y$ (where X and Y are the corresponding $m \times 1$ matrices).

Elementary matrices: These are $m \times m$ (square) matrices of one of the following three types:

$$S_{ij}^m := \mathbb{I}_m - \mathbb{E}_{ii}^m - \mathbb{E}_{jj}^m + \mathbb{E}_{ij}^m + \mathbb{E}_{ji}^m = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix},$$

$$S_i^m(\tfrac{1}{a}) := \mathbb{I}_m + (\tfrac{1}{a} - 1)\mathbb{E}_{ii}^m = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{1}{a} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

$$\text{and} \quad R_{ji}^m(-b) := \mathbb{I}_m - b\mathbb{E}_{ji}^m = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & -b & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

I'll drop the superscript m in the sequel sometimes.[3]

The elementary row operations of §I.B may be interpreted as left-multiplying the augmented matrix (representing our linear system) by one of these elementary matrices: the first exchanges the ith and j^{th} rows of the matrix it operates on; the second divides the ith row by a; and the third subtracts $b \times (i^{\text{th}}$ row$)$ from the jth row. All three types of matrices are clearly invertible, with

$$S_{ij}^{-1} = S_{ij}\,, \quad S_i(\tfrac{1}{a})^{-1} = S_i(a)\,, \quad \text{and} \;\; R_{ji}(-b)^{-1} = R_{ji}(b).$$

[3]The notation can be a bit confusing: $\mathbb{E}_{ij}$ is not the $(i,j)^{\text{th}}$ entry of something, it is a matrix. Its $(k,\ell)^{\text{th}}$ entry is $(\mathbb{E}_{ij})_{k\ell} = \delta_{ik}\delta_{j\ell}$. Similar remarks apply, for example, to S_{ij} and $R_{ij}(b)$.

Exercises

(1) Prove that $A\vec{x} = \sum_{j=1}^{n} x_j \vec{c}_j$ by calculating the ith entry of each.

(2) Find two different 2×2 matrices A such that $A^2 = 0$ but $A \neq 0$.

(3) Show that if square matrices A, B, and AB are symmetric, then $AB = BA$. Give an example with A and B symmetric, but AB not symmetric and $AB \neq BA$.

(4) Give a quick proof that columns of AB are linear combinations of columns of A. [Hint: what is $AB\hat{e}_j$?]

(5) By carrying out Gauss-Jordan and keeping track of your steps, find elementary matrices $E_1, \ldots, E_k$ such that $E_k \cdots E_2 E_1 A = \mathbb{I}$, where

$$A := \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 3 & 0 & 1 \end{pmatrix}.$$

(6) [Hamilton's quaternions] Consider the set $\mathbb{H} \subset M_2(\mathbb{C})$ (of 2×2 matrices with complex entries) of the form

$$x = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} a_0 + a_1\sqrt{-1} & a_2 + a_3\sqrt{-1} \\ -a_2 + a_3\sqrt{-1} & a_0 - a_1\sqrt{-1} \end{pmatrix}, \quad a_i \in \mathbb{R}.$$

Show that $\mathbb{H}$ is closed under addition and multiplication, and that every nonzero element is invertible. Give an example to show that multiplication is not commutative.