Basic Matrix Manipulations

The notes below contain simple ways to rewrite matrix-vector and matrix-matrix products that we will use repeatedly throughout the course.

Basic Notation

For an $M \times N$ matrix $\boldsymbol{A}$, we denote the columns as $\boldsymbol{a}_{c1}, \ldots, \boldsymbol{a}_{cN} \in \mathbb{R}^M$ and the rows as $\boldsymbol{a}_{r1}, \ldots, \boldsymbol{a}_{rM} \in \mathbb{R}^N$, and so

$$\boldsymbol{A} = \begin{bmatrix} | & | & & | \\ \boldsymbol{a}_{c1} & \boldsymbol{a}_{c2} & \cdots & \boldsymbol{a}_{cN} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} - & \boldsymbol{a}_{r1}^{\mathrm{T}} & - \\ - & \boldsymbol{a}_{r2}^{\mathrm{T}} & - \\ & \vdots & \\ - & \boldsymbol{a}_{rM}^{\mathrm{T}} & - \end{bmatrix}.$$

We will often refer to the $\{\boldsymbol{a}_{rm}\}$ as "the rows of $\boldsymbol{A}$", even though, strictly speaking, the $\{\boldsymbol{a}_{rm}\}$ are column vectors in $\mathbb{R}^N$, and it is the $\{\boldsymbol{a}_{rm}^{\mathrm{T}}\}$ that are the $1 \times N$ rows of $\boldsymbol{A}$.

The entries of a vector in $\mathbb{R}^N$ and an $M \times N$ matrix will be denoted using brackets:

$$\boldsymbol{x} = \begin{bmatrix} x[1] \\ x[2] \\ \vdots \\ x[N] \end{bmatrix}, \qquad \boldsymbol{A} = \begin{bmatrix} A[1,1] & A[1,2] & \cdots & A[1,N] \\ A[2,1] & A[2,2] & \cdots & A[2,N] \\ \vdots & & \ddots & \\ A[M,1] & A[M,2] & \cdots & A[M,N] \end{bmatrix}.$$

Note that vectors $\boldsymbol{x}$ and matrices $\boldsymbol{A}$ are typeset in bold, while their entries $x[n]$ and $A[m,n]$ are not, since they are scalars.

Matrix-vector multiplies

We can think of the action of an $M \times N$ matrix $\boldsymbol{A}$ on a vector $\boldsymbol{x} \in \mathbb{R}^N$ in one of two ways.

The first is as a series of inner products against the rows of $\boldsymbol{A}$:

$$\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_{r1}^{\mathrm{T}}\boldsymbol{x} \\ \boldsymbol{a}_{r2}^{\mathrm{T}}\boldsymbol{x} \\ \vdots \\ \boldsymbol{a}_{rM}^{\mathrm{T}}\boldsymbol{x} \end{bmatrix}.$$

The other is as a linear combination of the columns of $\boldsymbol{A}$:

$$\boldsymbol{A}\boldsymbol{x} = \sum_{n=1}^{N} x[n]\boldsymbol{a}_{cn}.$$

Notes by J. Romberg and M. Davenport – August 18, 2019 – 21:05

Matrix-matrix multiplies

Likewise, the product of an $M \times N$ matrix $\boldsymbol{A}$ and a $N \times P$ matrix $\boldsymbol{B}$ can be thought of as a collection of the inner products between all of the rows of $\boldsymbol{A}$ and all of the columns of $\boldsymbol{B}$,

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_{r1}^{\mathrm{T}}\boldsymbol{b}_{c1} & \boldsymbol{a}_{r1}^{\mathrm{T}}\boldsymbol{b}_{c2} & \cdots & \boldsymbol{a}_{r1}^{\mathrm{T}}\boldsymbol{b}_{cP} \\ \boldsymbol{a}_{r2}^{\mathrm{T}}\boldsymbol{b}_{c1} & \boldsymbol{a}_{r2}^{\mathrm{T}}\boldsymbol{b}_{c2} & \cdots & \boldsymbol{a}_{r2}^{\mathrm{T}}\boldsymbol{b}_{cP} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{rM}^{\mathrm{T}}\boldsymbol{b}_{c1} & \boldsymbol{a}_{rM}^{\mathrm{T}}\boldsymbol{b}_{c2} & \cdots & \boldsymbol{a}_{rM}^{\mathrm{T}}\boldsymbol{b}_{cP} \end{bmatrix},$$

as a sum of the rank 1 matrices formed by taking the outer product of the columns of $\boldsymbol{A}$ with the rows of $\boldsymbol{B}$,

$$\boldsymbol{A}\boldsymbol{B} = \sum_{n=1}^{N} \boldsymbol{a}_{cn}\boldsymbol{b}_{rn}^{\mathrm{T}},$$

as left action of $\boldsymbol{A}$ on the collective columns of $\boldsymbol{B}$,

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} | & | & & | \\ \boldsymbol{A}\boldsymbol{b}_{c1} & \boldsymbol{A}\boldsymbol{b}_{c2} & \cdots & \boldsymbol{A}\boldsymbol{b}_{cP} \\ | & | & & | \end{bmatrix}$$

or as right action of $\boldsymbol{B}$ on the rows of $\boldsymbol{A}$

$$\boldsymbol{A}\boldsymbol{B} = \begin{bmatrix} - & \boldsymbol{a}_{r1}^{\mathrm{T}}\boldsymbol{B} & - \\ - & \boldsymbol{a}_{r2}^{\mathrm{T}}\boldsymbol{B} & - \\ & \vdots & \\ - & \boldsymbol{a}_{rM}^{\mathrm{T}}\boldsymbol{B} & - \end{bmatrix}.$$

We again stress that these are just four different ways to write down exactly the same thing.

Second-order forms

For an $N \times N$ matrix $\boldsymbol{A}$ and a vector $\boldsymbol{x} \in \mathbb{R}^N$, the quadratic form $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ can be expanded as

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \sum_{m=1}^{N}\sum_{n=1}^{N} A[m,n]x[m]x[n].$$

Similarly, for an $M \times N$ matrix $\boldsymbol{A}$ and vectors $\boldsymbol{y} \in \mathbb{R}^M,\ \boldsymbol{x} \in \mathbb{R}^N$, the bilinear form $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ can be expanded as

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \sum_{m=1}^{M}\sum_{n=1}^{N} A[m,n]y[m]x[n].$$

Note that if $\boldsymbol{D}$ is an $N \times N$ diagonal matrix, so $D[m,n] = 0$ for $m \neq n$, then

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x} = \sum_{n=1}^{N} D[n,n]x[n]^2.$$

Three matrices

Let $\boldsymbol{U}$ be an $M \times N$ matrix, $\boldsymbol{C}$ a $N \times P$ matrix, and $\boldsymbol{W}$ a $P \times Q$ matrix. Then the $M \times Q$ matrix $\boldsymbol{UCW}$ can be written as

$$\boldsymbol{UCW} = \begin{bmatrix} \boldsymbol{u}_{r1}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c1} & \boldsymbol{u}_{r1}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c2} & \cdots & \boldsymbol{u}_{r1}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{cQ} \\ \boldsymbol{u}_{r2}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c1} & \boldsymbol{u}_{r2}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c2} & \cdots & \boldsymbol{u}_{r2}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{cQ} \\ \vdots & & \ddots & \\ \boldsymbol{u}_{rM}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c1} & \boldsymbol{u}_{rM}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{c2} & \cdots & \boldsymbol{u}_{rM}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{w}_{cQ} \end{bmatrix},$$

or

$$\boldsymbol{UCW} = \sum_{n=1}^{N}\sum_{p=1}^{P} C[n,p]\,\boldsymbol{u}_{cn}\boldsymbol{w}_{rp}^{\mathrm{T}}.$$

In the special case where $\boldsymbol{C}$ is square and diagonal

$$\boldsymbol{C} = \begin{bmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_N \end{bmatrix},$$

then the above reduces to

$$\boldsymbol{UCW} = \sum_{n=1}^{N} c_n \boldsymbol{u}_{cn}\boldsymbol{w}_{rn}^{\mathrm{T}}.$$