Matrices and Vectors

Notation:

a represents a scalar, ie, a number (real or complex).

$\underline{a}$ represents a column vector.

$\underline{\underline{A}}$ represents a matrix.

Suppose we are dealing with a 3-dimensional space, then:

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

$$\underline{\underline{A}} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

The transpose of vector $\underline{a}$, which is denoted by $\underline{a}^T$, is a row vector with the same components as $\underline{a}$. By taking the transpose of $\underline{a}$, we transform the single column of $\underline{a}$ into a single row. Likewise, taking the transpose of a matrix $\underline{\underline{A}}$ consists of changing every column of $\underline{\underline{A}}$ into a row in $\underline{\underline{A}}^T$ (or, changing every row of $\underline{\underline{A}}$ into a column in $\underline{\underline{A}}^T$). So, using the same $\underline{a}$ and $\underline{\underline{A}}$ as above, we have this transpose vector and transpose matrix:

$$\underline{a}^T = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}$$

$$\underline{\underline{A}}^T = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$$

or, to take a specific example:

$$\text{If} \qquad \underline{\underline{B}} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$\text{then} \qquad \underline{\underline{B}}^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$$

Another way to express the relation between a matrix $\underline{\underline{A}}$ and its transpose $\underline{\underline{A}}^T$ is to simply indicate what happens to an arbitrary element A_{ij} of the matrix.

$$(\underline{\underline{A}}^T)_{ij} = A_{ji} \qquad \text{for all i, j}$$

The above equation can be described in words like this: "*the element ij of the matrix* $\underline{\underline{A}}^T$) *is the element ji of the matrix* $\underline{\underline{A}}$." The product of a vector $\underline{x}$ with a number a, and product of a matrix $\underline{\underline{X}}$ with a number a are defined by these relations:

$$a\underline{x} = a \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ ax_3 \end{pmatrix}$$

$$a\underline{\underline{X}} = a \begin{pmatrix} X_{11} & X_{21} & X_{31} \\ X_{12} & X_{22} & X_{32} \\ X_{13} & X_{23} & X_{33} \end{pmatrix} = \begin{pmatrix} aX_{11} & aX_{21} & aX_{31} \\ aX_{12} & aX_{22} & aX_{32} \\ aX_{13} & aX_{23} & aX_{33} \end{pmatrix}$$

Here again we could save a lot of writing by simply stating what happens to an arbitrary component of a vector, or an arbitrary element of a matrix, upon multiplication by a. We write

$$(a\underline{x})_j = ax_j$$

$$(a\underline{\underline{X}})_{ij} = a\underline{\underline{X}}_{ij}$$

The word description of these two equations would be:
"*The j'th component of the vector obtained by multiplying the vector* $\underline{x}$ *by the number a is* ax_j";
and
"*The element ij of the matrix obtained by multiplying the matrix* $\underline{\underline{X}}$ *by the number a is* aX_{ij}".

Scalar Products

The scalar product of a N-dimensional vector $\underline{y}$ by a N-dimensional vector $\underline{x}$, denoted $\underline{x}^T \cdot \underline{y}$ or simply $\underline{x}^T \underline{y}$, is defined by

$$\underline{x}^T \cdot \underline{y} \ = \ \sum_{j=1}^{N} x_j y_j$$

Note that the scalar product of two vectors is a scalar (a number). Another name for scalar product is "inner product". If we write things in full,

$$\underline{x}^T \cdot \underline{y} \ = \ \begin{pmatrix} x_1 & x_2 & x_3 & \ldots & x_N \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \ldots \\ y_N \end{pmatrix}$$

$$= x_1 y_1 + x_2 y_2 + x_3 y_3 + \ldots + x_N y_N$$

The inner product of a vector $\underline{x}$ by a matrix $\underline{\underline{A}}$ is defined by

$$\underline{y} \ = \ \underline{\underline{A}}\, \underline{x}$$

$$y_i \ = \ \sum_{j=1}^{N} A_{ij} x_j$$

The inner product of a matrix $\underline{\underline{A}}$ and a matrix $\underline{\underline{B}}$ is defined by

$$\underline{\underline{P}} = \underline{\underline{A}}\,\underline{\underline{B}}$$
$$P_{ik} = \sum_{j=1}^{N} A_{ij} B_{jk}$$

Generally speaking

$$\underline{\underline{A}}\,\underline{\underline{B}} \neq \underline{\underline{B}}\,\underline{\underline{A}}$$

It may happen, *sometimes*, that $\underline{\underline{A}} \;\; \underline{\underline{B}} = \underline{\underline{B}} \;\; \underline{\underline{A}}$. In such a case, it is said that the matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$ *commute*.

The **trace of a matrix** is the sum of its diagonal elements.

$$Tr(\underline{\underline{A}}) = \sum_{j=1}^{N} A_{jj}$$

Note that:

$$Tr(\underline{\underline{A}}\,\underline{\underline{B}}) = Tr(\underline{\underline{B}}\,\underline{\underline{A}})$$

even if $\underline{\underline{A}}\,\underline{\underline{B}} \neq \underline{\underline{B}}\,\underline{\underline{A}}$.

The **unit matrix**, denoted by $\underline{\underline{1}}$ or by $\underline{\underline{\mathrm{I}}}$, is a diagonal matrix with one's on the diagonal and zero's everywhere else.

$$(\underline{\underline{1}})_{ij} = 0 \quad \text{if } i \neq j$$

$$(\underline{\underline{1}})_{ii} = 1$$

$$\underline{\underline{1}}\,\underline{\underline{M}} = \underline{\underline{M}}\,\underline{\underline{1}} = \underline{\underline{M}}$$

where $\underline{\underline{M}}$ can be any N-by-N matrix.

Inverse of a matrix $\underline{\underline{\mathrm{A}}}$

The inverse of $\underline{\underline{\mathrm{A}}}$, denoted $\underline{\underline{\mathrm{A}}}^{-1}$, is defined by the equations

$$\underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{A}}^{-1}\,\underline{\underline{A}} = \underline{\underline{1}}$$

In general, it is not easy to find $\underline{\underline{\mathrm{A}}}^{-1}$ corresponding to a given matrix $\underline{\underline{\mathrm{A}}}$. But here are two simple examples.

$$\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2}{5} & \frac{-1}{10} \\ \frac{-1}{5} & \frac{3}{10} \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & \frac{1}{5} \end{pmatrix}$$

The second example is particularly simple: the inverse of a diagonal matrix $\underline{\underline{D}}$ is a diagonal matrix with diagonal elements equal to $1/D_{jj}$.

Not all matrices have an inverse. Here are three examples of matrices that do not have an inverse.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \ , \ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \ , \text{ and } \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

Set of basis vectors, or “basis”

Here’s an example in 3D of what we mean by **basis**. In 3D, it is a set of 3 vectors that satisfy certain conditions. Normally, the 3 vectors are (i) orthogonal, and (ii) normalized. An example:

$$\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$$

$$\underline{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Any 3D vector $\underline{a}$ can be written exactly as a *linear combination* of the three basis vectors:

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1 \, \underline{u}_1 + a_2 \, \underline{u}_2 + a_3 \, \underline{u}_3$$

$$= a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ a_3 \end{pmatrix}$$

There are many other possible choices of bases, for example $\{\ \underline{v}_1, \underline{v}_2, \underline{v}_3\ \}$ with

$$\underline{v}_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{pmatrix}\ ,\ \underline{v}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}\ ,\ \underline{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Any vector $\underline{a}$ could also be expanded in this new basis

$$\underline{a} = \sum_{j=1}^{3} c_j\ \underline{v}_j$$

How do we get the c_j's? Here's the general rule, which works for any vector $\underline{a}$ and any basis $\{\ \underline{v}_1, \underline{v}_2, \underline{v}_3\ \}$:

$$c_k = \underline{a}^T \cdot \underline{v}_k = \underline{v}_k^T \cdot \underline{a}$$

Incomplete basis

Suppose we have a 3D vector but only 2 basis vectors to describe it. Take an example:

$$\underline{x} = \begin{pmatrix} 0.8651 \\ 0.5000 \\ 0.0400 \end{pmatrix}$$

$$\text{with basis } \underline{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ \underline{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Note that $\underline{x}^T \cdot \underline{x} = 1$. What is the best approximation to $\underline{x}$ that we can make with this incomplete basis? We require that the approximate vector $\underline{y} \approx \underline{x}$ is normalized, like $\underline{x}$.

$$\underline{y} = \left[(\underline{x}^T \cdot \underline{u}_1)\, \underline{u}_1 + (\underline{x}^T \cdot \underline{u}_2)\, \underline{u}_2 \right] \times N$$

N is a normalization factor that will be chosen to make $\underline{y}^T \cdot \underline{y} = 1$. With these formulas, you will find that $c_1 = 0.8651$, $c_2 = 0.5000$, $N = 1.000802$, and

$$\underline{y}^T \cdot \underline{x} = 0.999198$$

The scalar product being very close to 1 means that the vectors $\underline{y}$ and $\underline{x}$ are very similar (the scalar product would be 1 only if the two vectors were identical). Now let's try with a different basis:

$$\underline{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \ \underline{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\underline{y} \ = \ \left[(\underline{x}^T \cdot \underline{u}_1) \ \underline{u}_1 + (\underline{x}^T \cdot \underline{u}_3) \ \underline{u}_3 \right] \times N$$

This time we find $c_1 = 0.8651$, $c_3 = 0.04$, $N = 1.15470$, and

$$\underline{y}^T \cdot \underline{x} \ = \ 0.8660$$

This is a lot worse than before. The reason for this is clear. The component x_2 is much greater than x_3. So, for the purpose of describing $\underline{x}$, a basis that can not describe the second component of $\underline{x}$ is a lot worse than a basis that can not describe the third component of $\underline{x}$.

Eigenvectors of matrices

By definition an eigenvector $\underline{v}$ associated to a matrix $\underline{\underline{\mathrm{A}}}$ satisfied the eigenvalue equation

$$\underline{\underline{A}}\;\underline{v} = \lambda\;\underline{v}$$

Where λ is a number (real or complex). If $\underline{\underline{A}}$ is a N-by-N matrix, it has N eigenvalue-eigenvector pairs $(\lambda_i, \underline{v}_i)$. Note that if $\underline{v}_i$ is an eigenvector of $\underline{\underline{A}}$ with eigenvalue λ_i, than $c\underline{v}_i$ is also an eigenvector of $\underline{\underline{A}}$ with eigenvalue λ_i, no matter what c is. That's easy to show:

$$\begin{aligned}
\underline{\underline{A}}\;(c\underline{v}_i) &= \lambda_i\;(c\underline{v}_i) \\
c\;\underline{\underline{A}}\;\underline{v}_i &= c\;\lambda_i\;\underline{v}_i \\
\underline{\underline{A}}\;\underline{v}_i &= \lambda_i\;\underline{v}_i
\end{aligned}$$

Here's an example of a 4-by-4 matrix and its 4 eigenvalue-eigenvector pairs.

$$\underline{\underline{A}} = \begin{pmatrix} 1 & b & 0 & b \\ b & 1 & b & 0 \\ 0 & b & 1 & b \\ b & 0 & b & 1 \end{pmatrix}$$

You can verify for yourself that the following are eigenvalues and eigenvectors of $\underline{\underline{A}}$

$$\lambda_1 = (1+2b)\ ,\ \underline{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 1\ ,\ \underline{v}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

$$\lambda_3 = 1\ ,\ \underline{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$\lambda_4 = (1-2b)\ ,\ \underline{v}_4 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

Note that: (i) these eigenvectors are mutually orthogonal; (ii) the eigenvectors are not normalized, but we only need to multiply them by $\frac{1}{2}$ to get normalized eigenvectors (eigenvalues are unchanged); (iii) these $\underline{v}_i$'s form a complete basis for expanding any 4-dimensional vector.