Matrix Algebra Summary

Konrad Menzel

October 16, 2025

For our purposes, matrices are a notational "trick" to organize a large number of variables.

An $n \times k$ **matrix A** is an array of n rows and k columns that has the form

$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

By convention, the first subscript of a typical element of the matrix denotes the row, and the second subscript the column in which it is located.

The **transpose** of A is given by

$$\mathbf{A}' = [a_{ji}] = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & & \ddots & \vdots \\ a_{1k} & & \cdots & a_{nk} \end{bmatrix}$$

so that the transpose of an $n \times k$ matrix is a $k \times n$ matrix. The lth column of A becomes the lth row of the transpose A', and the mth row of A becomes the mth column of A'.

A k-dimensional **row vector** v is an $1 \times k$ matrix

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_k \end{bmatrix}$$

and an n-dimensional **column vector** u is an $m \times 1$ matrix

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

We usually use uppercase letters to denote general matrices, and lowercase letters to denote vectors.

A $k \times k$ matrix - i.e. a matrix that has the same number of rows and columns - is called a **square matrix**.

The $n \times k$ **null matrix 0** is a matrix with n rows and k columns whose elements are all equal to zero.

Matrix addition: If we have two $n \times k$ matrices $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$, then we can add the two matrices by adding their respective entries element by element

$$\mathbf{A} + \mathbf{B} = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}] = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1k} + b_{1k} \\ a_{21} + b_{21} & a_{22} + b_{22} & & \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nk} + b_{nk} \end{bmatrix}$$

The **matrix sum** of **A** and **B** is again an $n \times k$ matrix. Note that it is crucial that **A** and **B** have the same dimension.

Scalar multiplication: If $c \in \mathbb{R}$ is a real number (**scalar**), then

$$c\mathbf{A} = [ca_{ij}] = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1k} \\ ca_{21} & ca_{22} & & \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nk} \end{bmatrix}$$

i.e. each element of **A** is multiplied with the same constant c.

The **matrix product** is only defined between matrices **A** and **B** that are **conformable** in the following sense: For the matrix product $\mathbf{A} \cdot \mathbf{B}$ to be defined, the number of columns of **A** has to be equal to the number of rows of **B**. For an $m \times n$ matrix $\mathbf{A} = [a_{ij}]$ and an $n \times k$ matrix $\mathbf{B} = [b_{ij}]$, the matrix product $\mathbf{A} \cdot \mathbf{B}$ is an $m \times k$ matrix of the form

$$\underset{(m \times n)}{\mathbf{A}} \quad \cdot \quad \underset{(n \times k)}{\mathbf{B}} \quad = \quad \left[\textstyle\sum_{l=1}^{n} a_{il} b_{lj}\right] \quad =: \quad \underset{(m \times k)}{\mathbf{C}}$$

It is therefore always very important to keep track of the dimensions of matrices to make sure that the product is in fact defined.

The **inner product** of two $n \times 1$ (column) vectors **u** and **v** is defined as the matrix product

$$\mathbf{u}' \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i$$

In particular, since $\mathbf{u}'$ is an $1 \times n$ matrix and **v** is an $n \times 1$ matrix, the inner product is an 1×1 matrix, which is just a real number.

The k-dimensional **identity matrix** $\mathbf{I}_k$ is a $k \times k$ square matrix whose diagonal elements are all equal to 1, and whose off-diagonal elements all equal zero:

$$\mathbf{I}_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

We can easily check that for an $n \times k$ matrix $\mathbf{A}$ we have

$$\mathbf{I_n} \cdot \mathbf{A} = \mathbf{A} \quad \text{and } \mathbf{A} \cdot \mathbf{I_k} = \mathbf{A}$$

A square $k \times k$ matrix $\mathbf{A}$ may have an **inverse matrix** $\mathbf{A}^{-1}$ such that the produc

$$\mathbf{A} \cdot \mathbf{A^{-1}} = \mathbf{A^{-1}} \cdot \mathbf{A} = \mathbf{I_k}$$

Of course $\mathbf{A}^{-1}$ also has to be a $k \times k$ matrix for these products to be defined. A matrix $\mathbf{A}$ is only invertible if its columns are linearly independent vectors, i.e. we can't find a linear combination $z_1[a_{i1}] + z_2[a_{i2}] + \ldots z_k[a_{ik}]$ of the column vectors $\mathbf{a}_j$ of $\mathbf{A}$ that is equal to zero unless all the coefficients of that linear combination are zero, $z_1 = z_2 = \cdots = z_k = 0$. We then also say that $\mathbf{A}$ has full column rank.

We can also work with **random matrices**. Let

$$\mathbf{U} = \begin{bmatrix} U_{11} & U_{12} & \cdots & U_{1k} \\ U_{21} & U_{22} & & \\ \vdots & \vdots & \ddots & \vdots \\ U_{n1} & U_{n2} & \cdots & U_{nk} \end{bmatrix}$$

be a matrix whose elements are all random variables. Then the **expectation** of the random vector is a vector containing the expectation of each of the elements of $\mathbf{v}$:

$$\mathbb{E}[\mathbf{U}] = \begin{bmatrix} \mathbb{E}[U_{11}] & \mathbb{E}[U_{12}] & \cdots & \mathbb{E}[U_{1k}] \\ \mathbb{E}[U_{21}] & \mathbb{E}[U_{22}] & & \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}[U_{n1}] & \mathbb{E}[U_{n2}] & \cdots & \mathbb{E}[U_{nk}] \end{bmatrix}$$

The **variance-covariance matrix** of a k-dimensional random vector $\mathbf{v} := [V_1, \ldots, V_k]$ is defined as

$$\text{Var}(\mathbf{v}) := \mathbb{E}[(\mathbf{v} - \mathbb{E}[\mathbf{v}])(\mathbf{v} - \mathbb{E}[\mathbf{v}])'] = \begin{bmatrix} \text{Var}(V_1) & \text{Cov}(V_1, V_2) & \cdots & \text{Cov}(V_1, V_k) \\ \text{Cov}(V_1, V_2) & \text{Var}(V_2) & & \\ \vdots & \vdots & \ddots & \vdots \\ cov(V_1, V_k) & \text{Cov}(V_2, V_k) & \cdots & \text{Var}(V_k) \end{bmatrix}$$

The variance-covariance matrix of a k-dimensional random vector is a $k \times k$ matrix. Finally if $\mathbf{H}$ is a nonstochastic $m \times k$ matrix, then the expectation and variance of the $k \times 1$ random vector $\mathbf{v}$ satisfy

$$\mathbb{E}[\mathbf{H} \cdot \mathbf{v}] = \mathbf{H} \cdot \mathbb{E}[\mathbf{v}] \quad \text{and } \mathrm{Var}(\mathbf{H} \cdot \mathbf{v}) = \mathbf{H}\mathrm{Var}(\mathbf{v})\mathbf{H}'$$

In particular the variance of $\mathbf{H} \cdot \mathbf{v}$ is an $m \times m$ matrix.