An Introduction to Vectors and Matrices
Ed Stanek

Introduction

We introduce vectors and matrices for use in statistics. The introduction is intended for an audience with no previous matrix background. Basic terms and operations are defined and illustrated, along with the usual conventions for vector and matrix representation. In this introduction, we define the terms:

- scalar
- vector
- dimension
- transpose
- matrix
- conform;

define the operators:
- equality (of vectors and matrices)
- addition/subtraction
- multiplication by a scalar (of vectors and matrices)
- multiplication (vector by vector [inner product], vector by matrix, matrix by vector, matrix by matrix)

define the special vectors and matrices:
- vector of ones, $\mathbf{1}_n$
- square matrix of ones, $\mathbf{J}_n$
- identity matrix, $\mathbf{I}_n$

define the statistical representations:
- the design matrix
- the variance-covariance matrix.

Following this introduction, we illustrate how vectors and matrices can be used in simple statistical applications.

Basic Definitions and Notation for Vectors and Matrices

We define vectors, matrices and some terms and simple operations used with vectors and matrices. Each definition and operation is accompanied by a simple example.

Vectors:

A vector is an ordered set of numbers (or elements). The elements of a vector are called scalars. The dimension of a vector is the number of rows in the vector followed by the number of columns of the vector. Typically, vectors are only defined as a single column. Lower-case bold letters are used to represent vectors. The elements in the vector are represented by un-
bolded letters, or numbers. A vector and its elements can be represented compactly by
\( \mathbf{c} = (c_i) \), where the six elements in \( \mathbf{c} \) are the scalars \( c_i \) for \( i = 1, \ldots, 6 \).

**Example:**

\[
\mathbf{a} = \begin{pmatrix}
a_1 \\
a_2 \\
a_3 \\
a_4
\end{pmatrix} = \begin{pmatrix}
6 \\
3 \\
5 \\
-2
\end{pmatrix} \quad \text{or} \quad \mathbf{w} = \begin{pmatrix}
1 \\
0 \\
0
\end{pmatrix}.
\]

**Special Vectors:**

A vector in which all elements are equal to one occurs often in statistics. This vector is given a special representation as \( \mathbf{1}_n \), where \( n \) indicates the number of elements in the vector.

**Example:**

If \( \mathbf{v} = \begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
1
\end{pmatrix} \), then \( \mathbf{v} = \mathbf{1}_5 \).

**Transpose of a vector:**

The transpose of a vector (denoted by \( \mathbf{v}' \)) indicates that the column of a vector should be expressed as a row vector, with elements in the same order in the row as they were in the column.

**Example:**

If \( \mathbf{v} = \begin{pmatrix}
v_1 \\
v_2 \\
v_3
\end{pmatrix} \), then \( \mathbf{v}' = (v_1 \ v_2 \ v_3) \).

**Equality of Vectors:**

Vectors are equal only if they have the same dimensions (they conform) and the corresponding elements are equal.

**Example 1:** Let \( \mathbf{a} = \begin{pmatrix}
a_1 \\
a_2
\end{pmatrix} \) and \( \mathbf{b} = \begin{pmatrix}
b_1 \\
b_2
\end{pmatrix} \). Then \( \mathbf{a} = \mathbf{b} \) if and only if \( a_1 = b_1 \) and \( a_2 = b_2 \).
Example 2: If \( c' = (4 \, 0 \, -8.3) \) and \( d' = (4 \, 0 \, -8.3) \) then \( c = d \).

**Addition and Subtraction of Vectors:**

Vectors can be added and/or subtracted only if they have the same dimensions. In such cases, the vectors are said to conform for addition or subtraction. The addition (or subtraction) is formed by adding (or subtracting) corresponding elements of the vectors.

**Example:**

Let \( a = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \), \( b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \), and \( c = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \). Then

\[
\begin{align*}
a + b &= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}. \\
\end{align*}
\]

Also,

\[
\begin{align*}
a + b - c &= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.
\end{align*}
\]

The vector \( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \) is called a null vector.

**Multiplication of a vector by a scalar:**

Multiplication of a vector \( c = (c_i) \) by a scalar \( k \) involves multiplying each element of the vector, \( c_i \), by \( k \).

**Example:**

Suppose \( c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \) and \( k \) is a scalar. Then \( k c = \begin{pmatrix} k c_1 \\ k c_2 \\ k c_3 \end{pmatrix} \).

If \( c_1 = 2, \ c_2 = 4, \ c_3 = -1 \) and \( k = 2 \), then \( k c = 2 \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \\ -2 \end{pmatrix} \).

**Multiplication of Vectors** (called the inner product of two vectors):

Two vectors can be multiplied only when the number of columns of the first equals the number of rows of the second.
number of rows of the second. When this is true, the vectors conform for multiplication. An alternative representation of scalar multiplication is using summation notation.

Example 1:

If \( \mathbf{a}' = (a_1, a_2, a_3) \) and \( \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \), then

\[
\mathbf{a}' \mathbf{b} = (a_1, a_2, a_3) \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 = \sum_{i=1}^{3} a_i b_i .
\]

Example 2:

Suppose \( \mathbf{a} = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 3 \end{pmatrix} \) and \( \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 7 \\ 4 \end{pmatrix} \). Then

\[
\mathbf{a}' \mathbf{b} = (0, 2, 7, 4) \begin{pmatrix} 1 \\ 4 \\ 0 \\ 3 \end{pmatrix} = 0 + 8 + 0 + 12 = 20 .
\]

Transpose of the Product of Vectors:

The transpose of the product of two or more vectors is formed by multiplying the transposed vectors in reversed order.

Example 1: Let \( \mathbf{a}' = (a_1, a_2, a_3) \) and \( \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \), so that \( \mathbf{a}' \mathbf{b} = \sum_{i=1}^{3} a_i b_i . \) Then

\[
(\mathbf{a}' \mathbf{b})' = (b_1, b_2, b_3)' = \mathbf{b}' \mathbf{a} = \sum_{i=1}^{3} a_i b_i .
\]

Example 2: Let \( \mathbf{a}' = (a_1, a_2, a_3) \) \( \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \), and \( \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \). Then

\[
(\mathbf{a}' \mathbf{b} \mathbf{c})' = (c_1, c_2)' = \mathbf{c}' \mathbf{a} = \left( c_1 \sum_{i=1}^{3} a_i b_i , c_2 \sum_{i=1}^{3} a_i b_i \right) .
\]

Matrices:

A matrix is an array of elements in rows and columns. Bold capital letters are used to
represent matrices. For example, the $r \times c$ matrix $\mathbf{A}$ is represented by

$$
\mathbf{A}_{r \times c} = \begin{pmatrix}
    a_{11} & a_{12} & a_{13} & \cdots & a_{1c} \\
    a_{21} & a_{22} & a_{23} & \cdots & a_{2c} \\
    a_{31} & a_{32} & a_{33} & \cdots & a_{3c} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    a_{r1} & a_{r2} & a_{r3} & \cdots & a_{rc}
\end{pmatrix} = \left( (a_{ij}) \right).
$$

If the number of rows equals the number of columns, $r = c$, then the matrix is called a square matrix. Otherwise, the matrix is called a rectangular matrix.

**Example 1**: A square matrix: $\mathbf{A}_{2 \times 2} = \left( (a_{ij}) \right) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

**Example 2**: A rectangular matrix: $\mathbf{B}_{2 \times 3} = \left( (b_{ij}) \right) = \begin{pmatrix} 5 & 7 & 8 \\ 6 & -1 & 0 \end{pmatrix}$

**The transpose of a Matrix**: 

The transpose of a matrix is formed by interchanging the rows and columns.

**Example**: Let $\mathbf{A}_{2 \times 2} = \left( (a_{ij}) \right) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\mathbf{B}_{2 \times 3} = \begin{pmatrix} 5 & 7 & 8 \\ 6 & -1 & 0 \end{pmatrix}$. Then

$$
\mathbf{A}' = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 5 & 6 \\ 7 & -1 \\ 8 & 0 \end{pmatrix}.
$$

**Special Matrices: a Diagonal Matrix, the Identify Matrix, a Matrix of Ones**: 

Matrices that have certain special form have a special representation. A diagonal matrix, $\mathbf{D}_{v} = \left( (d_{ij}) \right)$, is a square matrix that has all off diagonal elements equal to zero (such that $d_{ij} = 0$ when $i \neq j$, and $d_{ii} = v_{i}$ where $\mathbf{v} = \left( (v_{i}) \right)$). An identity matrix, $\mathbf{I}_{n}$, is a diagonal matrix where $\mathbf{v} = \mathbf{I}_{n}$. A square matrix will all elements equal to one is represented as a matrix $\mathbf{J}_{n}$.

**Example 1**: Let $\mathbf{b}_{3 \times 1} = \begin{pmatrix} b_{1} \\ b_{2} \\ b_{3} \end{pmatrix}$. Then the diagonal matrix $\mathbf{D}_{b}$ is given by $\mathbf{D}_{b} = \begin{pmatrix} b_{1} & 0 & 0 \\ 0 & b_{2} & 0 \\ 0 & 0 & b_{3} \end{pmatrix}$.
Example 2: The identity matrix $I_4$ is given by $I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Example 3: The matrix of ones given by $J_3$ is given by $J_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Equality of Matrices: Matrices are equal only if they have the same dimensions and the corresponding elements are equal.

Example: Let $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix}$ and $D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \\ d_{31} & d_{32} \end{pmatrix}$. Then $C = D$ only if $c_{ij} = d_{ij}$ for all $i = 1, \ldots, 3; j = 1, 2$.

Addition and Subtraction of Matrices: Matrices can be added or subtracted only if they have the same dimensions (they conform for addition or subtraction). The addition or subtraction is formed by adding or subtracting corresponding elements.

Example: Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, and $C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Then $A + B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}$. Also $A + B - C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 2 \end{pmatrix}$.

Multiplication of a Matrix by a scalar: Multiplication of a matrix by a scalar involves multiplying each element of the matrix by the scalar.

Example: Let $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{pmatrix}$ and $k$ be a scalar. Then $kC = \begin{pmatrix} kc_{11} & kc_{12} \\ kc_{21} & kc_{22} \\ kc_{31} & kc_{32} \end{pmatrix}$.

Multiplication of Matrices and Vectors: Matrices and vectors can be multiplied together by expressing the matrix as a set of row or column vectors, and then multiplying the vectors. The number of columns of the first term in the multiplication must match (conform) the number of rows of the second term.
Example 1: Let \(A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \) and \(c' = (c_1 \quad c_2 \quad c_3)\). We can express \(A\) as two column vectors, \(A = (a_1 \quad a_2)\) where \(a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}\) and \(a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}\), or as three row vectors, such that \(A = (a_1 \quad a_2 \quad a_3)\) where \(a_i' = (a_{i1} \quad a_{i2})\).

Then \(c'A = c'(a_1 \quad a_2) = (c'a_1 \quad c'a_2) = \left(\sum_{i=1}^3 c_i a_{i1} \quad \sum_{i=1}^3 c_i a_{i2}\right)\).

Also, \(AB = a_1'b = a_2'b = a_3'b\) where \(a_i' = (a_{i1} \quad a_{i2})\).

Example 2: Let \(X = \begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}, \quad V = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \) and \(c' = (1 \quad -1)\).

Then \(Xb = \begin{pmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 17 \\ 26 \end{pmatrix}\). \(c'V = (1 \quad -1)\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} = (2 \quad -2),\)

\(c'Vc = (1 \quad -1)\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (2 \quad -2)\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 4.\)

**Multiplication of matrices:** Two matrices can be multiplied together only when the number of columns of the first matrix is equal to the number of rows of the second matrix. If this is true, then the matrices conform for multiplication. If we denote the product of two matrices as \(AB = C = \left(\begin{pmatrix} c_{ij} \end{pmatrix}\right)\), then the elements of \(C\) which are given by \(c_{ij}\) are the result of the vector
product of the $i^{th}$ row of $A$ with the $j^{th}$ column of $B$. In short, $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

**Example:** Let $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1(1) + 0(1) + 1(1) & 1(-1) + 0(-1) + 1(-1) \\ -1(1) + 2(1) + 3(1) & -1(-1) + 2(-1) + 3(-1) \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 4 & -4 \end{pmatrix}$.

**The Transpose of A Product of Matrices:**

The transpose of a product of matrices is formed by multiplying the transpose of each of the matrices in reverse order.

**Example 1.** The transpose of the product of matrices is given as: $(ABC)' = C'B'A'$

**Statistical Illustrations:**

**Example 1:** A matrix representation of a simple linear regression model.

A simple linear regression model is given by

$$\mu_{y|x_i} = \beta_0 + \beta_1 x_i$$

or equivalently for the $i = 1, ..., n$ subjects as

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

where $\beta_0$ is the intercept and $\beta_1$ is the slope. We can express this model simultaneously for all sample subjects using vectors and matrices. Define

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$ Then the simple linear regression model is given by

$$y = X\beta + \varepsilon.$$
Suppose data is available on three hospitals concerning their size (# of beds) and occupancy (%).

<table>
<thead>
<tr>
<th>Hospital (i)</th>
<th>Occupancy (%)</th>
<th>Number of beds</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>65</td>
<td>200</td>
</tr>
</tbody>
</table>

A scatter plot suggests a linear relationship between the occupancy rate and the number of beds. For another hospital, if the relationship holds, we may be able to predict the occupancy rate from the number of beds. The regression model is of the form,

\[ Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \]

where \( Y_i \) is the occupancy rate for the \( i^{th} \) selected hospital (a random variable) and \( y_i \) is the occupancy rate for hospital \( i \) with \( x_i \) beds (a realized value).

In matrix form,

\[
\begin{pmatrix}
Y_1 \\
y_2 \\
y_3
\end{pmatrix} = \begin{pmatrix}
1 & x_1 \\
1 & x_2 \\
1 & x_3
\end{pmatrix} \begin{pmatrix}
\beta_0 \\
\beta_1
\end{pmatrix} + \begin{pmatrix}
\epsilon_1 \\
\epsilon_2 \\
\epsilon_3
\end{pmatrix}
\]

or

\[
\begin{pmatrix}
40 \\
50 \\
65
\end{pmatrix} = \begin{pmatrix}
1 & 80 \\
1 & 120 \\
1 & 200
\end{pmatrix} \begin{pmatrix}
\beta_0 \\
\beta_1
\end{pmatrix} + \begin{pmatrix}
\epsilon_1 \\
\epsilon_2 \\
\epsilon_3
\end{pmatrix}.
\]

Terms in this model are referred to with special names. The vector \( Y \) is a vector of dependent random variables. The matrix \( X \) is a non-stochastic design matrix. The vector \( \beta \) is the parameter vector. The first column of the design matrix corresponds to the coefficients of the first parameter. The coefficients in the second column of the design matrix corresponds to the coefficients of the second parameter. It is sometimes helpful to write the design matrix as

\[
X = \begin{pmatrix}
1 & 80 \\
1 & 120 \\
1 & 200
\end{pmatrix}
\]

to remind us of the meaning of the columns.

**Example 3**: The Variance Matrix.

In many settings, there will be more than one parameter of interest. Estimates of the parameters can be summarized in a parameter vector, \( \hat{\beta}' = (\hat{\beta}_0 \quad \hat{\beta}_1 \quad \cdots \quad \hat{\beta}_2) \). The variance and covariances of these estimates are summarized in a variance matrix given by:
Example 4: Variance of a Linear Combination of Parameters

There will often be interest in estimating the mean and variance of a linear combination of parameter estimates, such as the difference between group means. The variance can be readily estimated using the definition of the linear combination. Let a linear combination of parameter estimates be defined as \( \mathbf{e}'\hat{\mathbf{\beta}} \) where \( \text{var}(\hat{\mathbf{\beta}}) = \mathbf{V} \). Then \( \text{var}(\mathbf{e}'\hat{\mathbf{\beta}}) = \mathbf{e}' \mathbf{V} \mathbf{e} \).