Basic matrix algebra and R

Frank Miller, Department of Computer and Information Science, Linköpings University

Fall 2023

Definition and notation

A matrix is a rectangular array of numbers. If matrix $\mathbf{A}$ has n rows and p columns, we say that it is a $n \times p$-matrix. For example, n observations and p variables gives an $n \times p$-matrix as follows:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

For example, the specific 3×2-matrix

$$\mathbf{X} = \begin{pmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 8 \end{pmatrix}$$

can be defined in R for example by

```
X <- matrix(c(1, 2,
              1, 5,
              1, 8), ncol=2, byrow=TRUE)
```

A column-vector is a matrix with only one column,

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

and a row-vector is a matrix with only one row,

$$\mathbf{b} = \begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix}.$$

You can define a vector in R with the c-function, e.g.:

```
y <- c(3, 4)
```

With this definition it is not yet specified if `y` is a column- or a row-vector, see below.

Elementary matrix operations

If

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{pmatrix},$$

then, their sum $\mathbf{S}$ is given by the elementwise sum:

$$\mathbf{S} = \mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1p}+b_{1p} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2p}+b_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{np}+b_{np} \end{pmatrix}.$$

In R:

```
S <- A+B
```

If $\mathbf{A}$ is a $n \times p$-matrix and $\mathbf{B}$ is a $p \times m$-matrix, their (matrix-)product is defined by

$$\mathbf{A}{*}\mathbf{B} = \begin{pmatrix} (a_{11} * b_{11} + a_{12} * b_{21} + \ldots + a_{1p} * b_{p1}) & \dots & (a_{11} * b_{1m} + a_{12} * b_{2m} + \ldots + a_{1p} * b_{pm}) \\ \vdots & & \vdots \\ (a_{n1} * b_{11} + a_{n2} * b_{21} + \ldots + a_{np} * b_{p1}) & \dots & (a_{n1} * b_{1m} + a_{n2} * b_{2m} + \ldots + a_{np} * b_{pm}) \end{pmatrix}.$$

In R:

```
S <- A %*% B
```

If $n = p = m$, `S <- A * B` would yield a result, too, but it gives the elementwise product, which usually is not desired and leads then to errors!

When you multiply a matrix with a vector defined with the c-function, the vector will be promoted to either a column- or row-vector (see R-documentation for `%*%`). In the following example using the 2×3-matrix `X` and the two-dimensional vector `y` as above, `y` is promoted to a column-vector, since the matrix multiplication is well-defined in this case:

```
> X %*% y

[1,]   11
[2,]   23
[3,]   35
```

Transposed matrix

If $\mathbf{X}$ is a $n \times p$-matrix, then the so called transposed matrix is the $p \times n$-matrix when interchanging rows and columns of $\mathbf{X}$. It is denoted by $\mathbf{X}^\top$. For example,

$$\mathbf{X} = \begin{pmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 8 \end{pmatrix}, \quad \mathbf{X}^\top = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 5 & 8 \end{pmatrix}.$$

For any given $n \times p$-matrix $\mathbf{X}$, the matrix $\mathbf{V} = \mathbf{X}^\top \mathbf{X}$ is a symmetric matrix (this means $\mathbf{V}^\top = \mathbf{V}$). For the above example,

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} 3 & 15 \\ 15 & 93 \end{pmatrix}.$$

Inverse matrix

If $\mathbf{A}$ is a $n \times n$-matrix, then the inverse matrix $\mathbf{A}^{-1}$ of $\mathbf{A}$ is the $n \times n$-matrix such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix,

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{pmatrix}.$$

For example, for the above matrix $\mathbf{V}$, the inverse can be computed as follows:

```
V    <- t(X) %*% X
Vinv <- solve(V)
Vinv
Vinv %*% V
```

We get here:

$$\mathbf{V}^{-1} = \begin{pmatrix} 1.7222 & -0.2778 \\ -0.2778 & 0.0556 \end{pmatrix} \quad \text{and} \quad \mathbf{V}^{-1}\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that computation of inverses of large matrices is computationally expensive. Note further that an alternative to `V <- t(X) %*% X` is `V <- crossprod(X)` which can be slightly faster.

Note that not every quadratic matrix has an inverse matrix. A matrix $\mathbf{A}$ which has an inverse is called invertible. If you use `solve` for a matrix which is not invertible, you will get an error message like the following:

```
> A <- matrix(c(1,0,0,0),ncol=2)
> solve(A)
Error in solve.default(A) :
  Lapack routine dgesv: system is exactly singular: U[2,2] = 0
```