Basics of linear algebra

christine.dillmann@universite-paris-saclay.fr

September 7th 2024

This memento on linear algebra is much inspired from the web page https://www.unilim.fr/pages_perso/jean.debord/math/matrices/matrices.htm. The credit is Jean Debord from Faculté de médecine of Limoges. I also used web ressources (including the valuable wikipaedia math pages). A special thank to Sandra Delaunay, Alexis Tchoudjem and Arnaud Bodin.

1 Definitions

A matrix $n \times m$ is a table with n lines and m columns containing scalars. Example with $n = 2$, $m = 3$:

$$A = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 3 \end{bmatrix}$$

n et m are the dimensions of the matrix. By convention, a_{ij} is the element of A contained in line i, column j. The A matrix can be described as:

$$A = [a_{ij}]_{n \times m}$$

1.1 Distinctive matrices

If $m = 1$, the matrix is a column vector:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

If $n = 1$, the matrix is a line vector and corresponds to the *transpotition* of the column vector $\mathbf{x}$:

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$$

If $n = m$, the matrix is square. Different square matrices can be created using R functions.

A *diagonal matrix* is a square matrix for which all non-diagonal elements are zero:

```r
diag(c(3,6,7))
```

```
##      [,1] [,2] [,3]
## [1,]    3    0    0
## [2,]    0    6    0
## [3,]    0    0    7
```

The *identity matrix*, called I_n is a $n \times n$ diagonal matrix for which diagonal elements are equal to 1:

```r
diag(nrow=4)
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1    0    0    0
## [2,]    0    1    0    0
## [3,]    0    0    1    0
## [4,]    0    0    0    1
```

A square matrix is *symmetrical* if for all $i \neq j$, $a_{ij} = a_{ji}$.

$$Z = \begin{bmatrix} 1 & 5 & 7 \\ 5 & 4 & 3 \\ 7 & 3 & 1 \end{bmatrix}$$

2 Matrix operations

Matrix operations are codified and need to be known.

2.1 Addition, substraction

Addition and substraction are performed term by term. The two matrices must have the same dimension.

$$A - \begin{bmatrix} 0 & 2 & 9 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 3 \end{bmatrix} - \begin{bmatrix} 0 & 2 & 9 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 3 & 2 \end{bmatrix}$$

2.2 Multiplication by a scalar

Each matrix term is multiplied by the scalar:

$$2 \times A = 2 \times \begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 10 & 14 \\ 4 & 8 & 6 \end{bmatrix}$$

2.3 Transposed

the transposed matrix A^T (can also be noted A') from a matrix A is obtained by exchanging rows and column of A. If

$$A = \begin{bmatrix} 1 & 5 & 7 \\ 2 & 4 & 3 \end{bmatrix}$$

then

$$A' = A^T = \begin{bmatrix} 1 & 2 \\ 5 & 4 \\ 7 & 3 \end{bmatrix}$$

The transposition of a column vector $(n \times 1)$ is a line vector $(1 \times n)$.

2.4 Matrices product

The *scalar product* of a row vector $\mathbf{x}^T_{1\times m}$ by a column vector $\mathbf{y}_{m\times 1}$ is a scalar obtained by summing the term by term products of each element of both vectors:

$$\mathbf{x}^T \times \mathbf{y} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \sum_{i=1}^{3} x_i \times y_i$$

Scalar product can only be computed between two vectors having the same dimensions.

The product of two matrices can only be computed if the column number of the left matrix is equal to the row number of the right matrix. The product of $A_{n\times m}$ by $B_{m\times p}$ is a matrix C with dimensions $n \times p$. The element c_{ij} is the scalar product of line i of matrix A by column j of matrix B:

$$c_{ij} = \sum_{k=1}^{m} a_{ik} \times b_{kj}$$

For example :

$$\begin{bmatrix}1 & 2\\5 & 4\\7 & 3\end{bmatrix} \times \begin{bmatrix}1 & 4 & 5 & 2\\2 & 2 & 3 & 4\end{bmatrix} = \begin{bmatrix}5 & 8 & 11 & 10\\13 & 28 & 37 & 26\\13 & 34 & 44 & 26\end{bmatrix}$$

In this example, $c_{11} = 1 + 2 \times 2 = 5$, et $c_{12} = 1 \times 4 + 2 \times 2 = 8$.

Matrices product has the following properties:

- *Associativity* : $A \times B \times C = (A \times B) \times C = A \times (B \times C)$
- *Distributivity* : $A \times (B + C) = A \times B + A \times C$
- *Non commutativity* : In general, $A \times B \neq B \times A$

The identity matrix is neutral with regard to multiplication:

$$A_{n\times m} \times I_m = I_n \times A_{n\times m} = A$$

Beware, the transposed matrix of a product is NOT the product of the transposed matrices: the order of the product changes:

$$(A \times B)' = B' \times A'$$

With R, the operator for matrix product is $\% * \%$:

```
A <- matrix(c(1,2,5,4,7,3),ncol=2,byrow=TRUE)
B <- matrix(c(1,4,5,2,2,2,3,4),ncol=4,byrow=TRUE)
A%*%B
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    5    8   11   10
## [2,]   13   28   37   26
## [3,]   13   34   44   26
```

Let's consider the column vector $\mathbf{x}_{n\times 1}$ and the square symmetrical matrix $A_{n\times n}$. Then

- The square root of the scalar product

$$\mathbf{x}' \times \mathbf{x} = \sum_{i=1}^{n} x_i^2$$

is called the *vector norm.* $||\mathbf{x}|| = \sqrt{\mathbf{x}' \times \mathbf{x}}$.

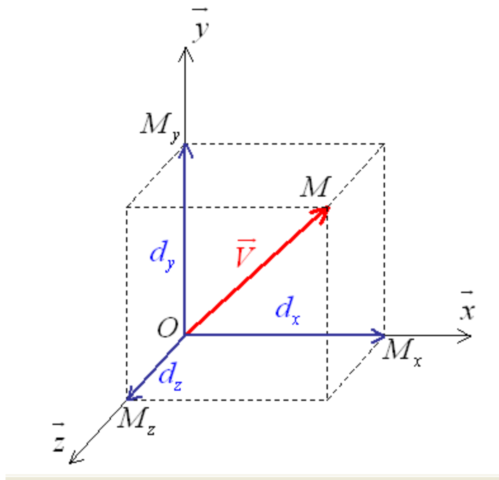
- Beware, the product $\mathbf{x} \times \mathbf{x}'$ is a symmetrical square matrix $n \times n$ that contains x_i^2 as diagonal terms, and the $x_i \times x_j$ products outside the diagonal.

- Le product $\mathbf{x}' \times A \times \mathbf{x}$ is called the *quadratic form.*

Exercise. Let $x = [2, 4, 3]$ and $A = \begin{bmatrix}1 & 3 & 5\\3 & 2 & 7\\5 & 7 & 2\end{bmatrix}$. Compute the norm of $\mathbf{x}$ and the quadratic form. Write the matrix $\mathbf{x} \times \mathbf{x}'$ and give its dimensions.*

2.5 Euclidian geometry

A column vector $\mathbf{x}_{n\times 1}$ represents the coordinates of a point in an euclidian space with orthonormal basis of dimension n. The norm of the vector is the distance between the point and the origin of the orthonormal basis.



3 Systems of linear equations

A system of n equations with n unkowns can be written in a matrix form. Let's consider the following system:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 & = & b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 & = & b_3 \end{array}$$

Such system also writes $A \times \mathbf{x} = b$ where $A = [a_{ij}]_{3,3}$ is a square matrix, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is a column vector (3×1) and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ is the column vector that contains the left members of the system.

Exercice. Write the A matrix and check that the product $A \times \mathbf{x}$ is a column vector. Compute its terms.

4 Properties of square matrices

Square matrices play an important role in linear algebra. We have already seen:

- The symmetrical matrix $\mathbf{x} \times \mathbf{x}'$ that describes some properties of a column vector.
- The A matrix that contains the linear coefficients of a system of equations.
- Note that whatever the dimensions of a given matrix $B_{n\times m}$, the matrix $C = B' \times B$ is a symmetrical matrix of dimensions $m \times m$

4.1 Square matrix inversion

A square matrix $A_{m\times m}$ is *inversible* or *regular* if there exist a square matrix A^{-1} (called *the inverse matrix*) such that :

$$A^{-1} \times A = A \times A^{-1} = I_m$$

If A^{-1} does not exist, the matrix is called *singular*.

Properties :

- $\left(A^{-1}\right)^{-1} = A$

- $(A')^{-1} = (A^{-1})'$
- $(A \times B)^{-1} = B^{-1} \times A^{-1}$ (beware the order of the product !)
- $Diag(D_{ii})^{-1} = 1/Diag(D_{ii})$

4.2 Determinant of a square matrix

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse matrix is :

$$A^{-1} = \frac{1}{ad - bc} \times \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where $ad - bc$ is called the *determinant* of the A matrix, noted $det(A)$.

Whatever the dimensions of a suqre matrix A, the inverse matrix A^{-1} exist only if $det(A) \neq 0$. If $det(A) = 0$, the A matrix is singular.

Properties of determinants :

- $det(A') = det(A)$
- $det(A \times B) = det(A) \times det(B)$
- the determinant of a diagonal or a triangular matrix is the product of its diagonal terms. In particular, $det(I_m) = 1$.

To compute the determinant of a matrix, one can use the Laplace method. If A_n is a square matrix of n columns, let $A_{i,j}$ the matrix of dimension $n-1$ obtained by removing the ith line and the jth column of A.

$$\det(A) = \sum_{j=1}^{n} a_{i;j}(-1)^{i+j}\det(A_{i,j}) = \sum_{i=1}^{n} a_{i;j}(-1)^{i+j}\det(A_{i,j})$$

$\det(A_{i,j})$ is called the minor of the term $a_{i,j}$.

Example :

$$\det(\begin{bmatrix} -2 & 2 & -3 \\ -1 & 1 & 3 \\ 4 & 0 & -1 \end{bmatrix}) = 2.(-1)^{(1+2)}.\det(\begin{bmatrix} -1 & 3 \\ 4 & -1 \end{bmatrix}) + 1.(-1)^{(2+2)}.\det(\begin{bmatrix} -2 & -3 \\ 4 & -1 \end{bmatrix})$$

4.3 Trace of a square matrix

The trace of a square matrix is the sum of its diagonal elements:

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

Two matrices sharing the same trace look alike.

4.4 Solving a system of linear equations

The solution of the system $A \times \mathbf{x} = \mathbf{b}$ depends on the characteristics of the A matrix:

4.4.1 If A_m is regular

Each member can be multiplied by A^{-1} :

$$A^{-1} \times A \times \mathbf{x} = A^{-1} \times \mathbf{b}$$

Using the properties above ($A^{-1} \times A = I_m$ et $I_m \times \mathbf{x} = \mathbf{x}$), we find

$$\mathbf{x} = A^{-1} \times \mathbf{b}$$

4.4.2 If A_m is of rank $r < m$

Imagine that p equations of the system are linear combinations of each other. For example in the following system $(n = 3)$:

$$\begin{aligned} x_1 + 3x_2 + x_3 &= 5 \\ 2x_1 + 6x_2 + 2x_3 &= 10 \\ 3x_1 + 9x_2 + 3x_3 &= 15 \end{aligned}$$

There are $p = 2$ equations that are linear combinations of the first equation. The system contains actually one equation with three unknowns. There is an infinity of possible solutions. The equation $x_1 + 3x_2 + x_3 = 5$ defines a curve (dimension 1) in the three dimentional space (x_1, x_2, x_3). The number $r = n - p = 1$ is called the *rank of the matrix*. The system is called *indeterminated*.

In that case, one can choose as a solution the point in the curve which is the closest to the origin, *i.e* the vector $\mathbf{x}$ that satisfies the constraint $x_1 + 3x_2 + x_3 = 5$ and which has the smallest norm.

4.4.3 Impossible system

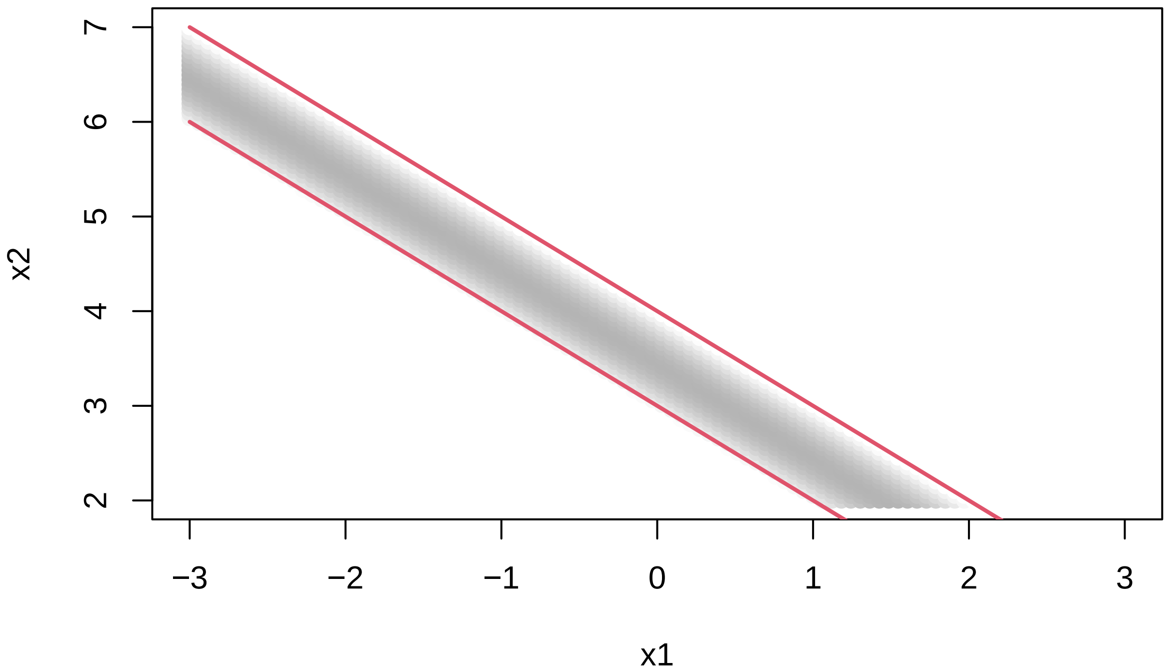
If the system's equations cannot be exprimed as linear combinations, the system has no solution. However, it is possible to compute the vector $\mathbf{x}$ such that the norm of $A \times \mathbf{x} - \mathbf{b}$ be minimal (although not equal to zero). Such vector is the best approximation of the solution in the sense of least squares.

Exercise. *We consider the following system*

$$\begin{aligned} x_1 + x_2 &= 3 \\ x_1 + x_2 &= 4 \end{aligned}$$

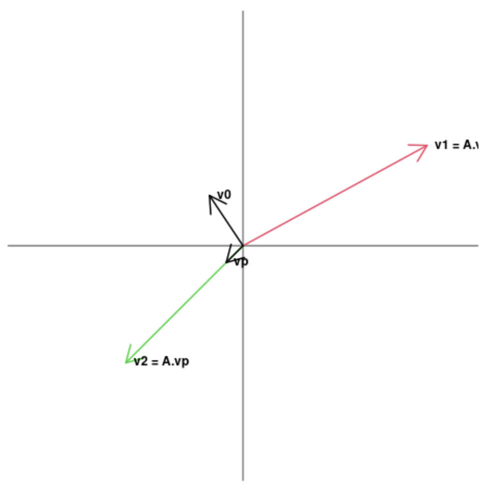
Interpret the results of the R script and the graphical output.

```
A <- matrix(c(1,1,1,1),ncol=2,byrow=TRUE)
b <-c(3,4)
n <- 100
x1 <- seq(-3,3,length=n)
x2 <- seq(2,7,length=n)
res <- NULL
for (i in 1:n){
  for (j in 1:n){
    z <- A%*%c(x1[i],x2[j]) - b
    dist <- sqrt(t(z)%*%z)
    res <- rbind.data.frame(res,cbind.data.frame(x1=x1[i],x2=x2[j],dist=dist))
  }
}
mycols <- res$dist
mycols[mycols>1] <- 1
plot(res$x1,res$x2,type="p",pch=19,col=grey(mycols),
     xlab="x1", ylab="x2")
points(res$x1, 4-res$x1,type="l", lwd=2,col=2)
points(res$x1, 3-res$x1,type = "l", lwd=2,col=2)
```



5 Eigenvalues, eigenvectors, and the characteristic polynomial of a matrix

A square matrix A $(n \times n)$ transforms a column vector x $(n \times 1)$ into a new vector $y = A.x$. In two dimensions, the figure below shows the transformation of the vector $v_0 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ by the matrix $A = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$, that leads to the vector v_1



Among all possible vectors that can be transformed by the A matrix some of them will only change the magnitude, but not the direction, as the vector v_p in the figure above.

Definition

- The scalar λ is an eigenvalue of a square matrix A is there exist a non-null vector v such that

$$A.v = \lambda.v$$

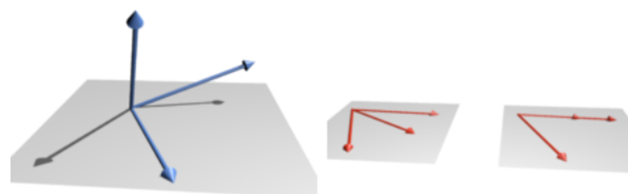
- The vector v is called the eigenvector of A, associated to the eigenvalue λ

5.1 Properties : distinct eigenvectors are linearly independent

A set of r vectors are linearly independent if the only linear combination of those vectors that is equal to the null vector is the one where all the coefficients are null :

$$\sum_{k=1}^{r} \alpha_k \cdot v_k = 0 <=> \alpha_1 = \alpha_2 = ... = 0$$

Independence means that none of the vectors from the family are linear combination one of the other. The maximum number of linearly independent vectors defines the vector space. In the figure below, the blue vectors from $\mathbb{R}^3$ are linearly independent. Considering the orange vectors, one of them is a linear combination of the two others. They are situated in the same plane.



Theorem : Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ **distinct** eigenvalues of A, and v_i the eigenvector associated to λ_i. Then, the vectors $v_1, v_2, \ldots, v_r$ are linearly independent.

The proof is the following. Suppose that $v_1, v_2, \ldots, v_{r-1}$ are linearly independent. $\sum_{k=1}^{r} \alpha_k \cdot v_k = 0$ means that $A.(\sum_{k=1}^{r} \alpha_k \cdot v_k) = \sum_{k=1}^{r} \alpha_k \cdot \lambda_k \cdot v_k = 0$, but also that $\lambda_1.(\sum_{k=1}^{r} \alpha_k \cdot v_k) = 0$.

Therefore, we need to show that

$$\sum_{k=1}^{r-1} \alpha_k \cdot (\lambda_k - \lambda_r) \cdot v_k = 0$$

Because $v_1, v_2, \ldots, v_{r-1}$ are linearly independent, we know that for all $k < r$, $\alpha_k \cdot (\lambda_k - \lambda_r) = 0$. We also know that the λ_k's are distinct, so $(\lambda_k - \lambda_r) \neq 0$. Therefore,

$$\alpha_1 = \alpha_2 = ... = \alpha_{r-1} = 0$$

The solution to $\sum_{k=1}^{r} \alpha_k \cdot v_k = 0$ is $\alpha_r \cdot v_r = 0$. Because v_r is an eigenvector of A, it cannot be the null vector. Therefore, we also have $\alpha_r = 0$.

5.2 Other matrices properties

Eigenvectors also have the following properties:

- If A_n is of full rank, there exist n non-zero eigenvalues. Otherwise, the number of non-zero eigenvalues is equal to the rank of the A matrix. The set the eigenvalues associated to the A matrix is called the matrix **spectrum**.
- $Tr(A) = \sum_{k=1}^{n} \lambda_k$
- $\det(A) = \prod_{k=1}^{n} \lambda_k$
- If one eigenvalue of A is null, than $\det(A) = 0$ and A is not inversible.
- If A is inversible, then the eigenvalues of A^{-1} are $\frac{1}{\lambda_k}$

5.3 Computing the eigenvalues of a matrix

$$A.v_k = \lambda_k.v_k <=> (A - \lambda_k.I).v_k = 0$$

If $(A - \lambda_k.I)$ was inversible, we would have $(A - \lambda_k.I)^{-1}(A - \lambda_k.I).v_k = v_k = 0$ which is not possible because an eigenvector cannot be the null vector. Therefore, $(A - \lambda_k.I)$ is not inversible, and

$$\det(A - \lambda.I) = 0$$

Finding the eigenvalues of a matrix is solving the equation above that is a polynomial of order n, called the characteristic polynomial.

Example: Let's consider the 3 matrix $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$. Using the properties of the determinant, it can be shown that

$$\det(A - \lambda.I) = 4 - 8 \cdot \lambda + 5 \cdot \lambda^2 - \lambda^3$$

Such a polynomial has three roots, 1, 2 and 2.

5.4 Eigenvectors

Given a particular eigenvalue λ_k of matrix A, one can define the set E_k of all the vectors v that satisfy the equation $(A - \lambda_k.I) \cdot v = 0$: $E_k = \{v; (A - \lambda_k.I) \cdot v = 0\}$

This set is the kernel or nullspace of the matrix $(A - \lambda_k.I)$. E_k is called the **eigenspace** or characteristic space of matrix A associated with λ_k.

Example: Using the same example for A as before, the eigenspace of A associated with $\lambda_1 = 1$ can be found y solving

$$(A - 1 \cdot I) \cdot v = \begin{bmatrix} 0 & 2 & 2 \\ 0 & 1 & 1 \\ -1 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

which gives the equations system

$$\begin{array}{rcl} y + z & = & 0 \\ -x + 2y + z & = & 0 \end{array}$$

or, alternatively,

$$\begin{array}{rcl} y & = & -z \\ x & = & z \end{array}$$

Here, the rank of the matrix $(A - \lambda_1 \cdot I)$ is two and the eigenspace contains any vector of the form $\begin{bmatrix} -z \\ z \\ z \end{bmatrix}$.

One can pick any value for $z \neq 0$ (remember, eigenvectors are not null). Hence, $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector associated with the eigenvalue $\lambda_1 = 1$ of the matrix A.

5.5 Matrix diagonalization and eigendecomposition

Let A be a square $n \times n$ matrix. If A has n linearly independent vectors q_i $(i \in 1..n)$, it can be factorized as

$$A = Q \Lambda Q^{-1}$$

where Q is the square $n \times n$ matrix whose ith columns is the eigenvector q_i of A, and Λ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues $\Lambda_{ii} = \lambda_i$.

Such a decomposition can be derived from the fundamental properties of the eigenvectors. Because $A \cdot q_i = \lambda_i \cdot q_i$, we also have (remember matrix multiplication properties) $A \cdot Q = Q \cdot \Lambda$. Multiplying by Q^{-1} gives $A = Q\Lambda Q^{-1}$

Similar matrices. In linear algebra, two $n \times n$ A and B matrices are similar if there exist a singular matrix P such that

$$A = P \cdot B \cdot P^{-1}$$

Two similar matrices represent the same linear transformation. They share the same rank, the same trace, the same determinant and the same eigenvalues.

Whenever a matrix A is similar to a diagonal matrix B, it is *diagonalizable*, that much simplifies algebraic computations. For example, the inverse of a diagonalizable matrix is

$$A^{-1} = P^{-1} \cdot B^{-1} \cdot P$$

Whenever all the eigenvalues of the A matrix are distinct, A is diagonalizable and the matrices A and Λ are similar.