MATRICES part 2

(modified content from Wikipedia articles on matrices http://en.wikipedia.org/wiki/Matrix_(mathematics))

3. Linear equations

Matrices can be used to compactly write and work with systems of linear equations. For example, if **A** is an *m-by-n* matrix, **x** designates a column vector (i.e., $n\times1$-matrix) of n variables $x_1, x_2, ..., x_n$, and **b** is an $m\times1$-column vector, then the matrix equation

$$\mathbf{Ax} = \mathbf{b}$$

is equivalent to the system of linear equations

$$a_{1,1}x_1 + a_{1,2}x_2 + ... + a_{1,n}x_n = b_1$$

$$...$$

$$a_{m,1}x_1 + a_{m,2}x_2 + ... + a_{m,n}x_n = b_m \ .$$

For example,

$$3x_1 + 2x_2 - x_3 = 1$$
$$2x_1 - 2x_2 + 4x_3 = -2$$
$$-x_1 + 1/2x_2 - x_3 = 0$$

is a system of three equations in the three variables x_1, x_2, and x_3. This can be written in the matrix form $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & -2 & 4 \\ -1 & 1/2 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

A **solution** to a linear system is an assignment of numbers to the variables such that all the equations are simultaneously satisfied. A solution to the system above is given by

$$x_1 = 1$$
$$x_2 = -2$$
$$x_3 = -2$$

since it makes all three equations valid.

A <mark>linear system may behave in any one of three possible ways</mark>:

1. The system has *infinitely many solutions*.
2. The system has a single *unique solution*.
3. The system has *no solution*.

3.1. Solving linear equations

There are several algorithms for solving a system of linear equations.

Elimination of variables

The <mark>simplest method</mark> for solving a system of linear equations is to repeatedly eliminate variables. This method can be described as follows:

1. In the first equation, <mark>solve for one of the variables in terms of the others</mark>.
2. <mark>Substitute this expression into the remaining equations</mark>. This yields a system of equations with one fewer equation and one fewer unknown.
3. Continue until you have reduced the system to a single linear equation.
4. Solve this equation, and then back-substitute until the entire solution is found.

For example, consider the following system:

$$x + 3y - 2z = 5$$
$$3x + 5y + 6z = 7$$
$$2x + 4y + 3z = 8$$

Solving the first equation for x gives $x = 5 + 2z - 3y$, and plugging this into the second and third equation yields

$$-4y + 12z = -8$$
$$-2y + 7z = -2$$

Solving the first of these equations for y yields $y = 2 + 3z$, and plugging this into the second equation yields $z = 2$. We now have:

$$x = 5 + 2z - 3y$$
$$y = 2 + 3z$$
$$z = 2$$

Substituting $z = 2$ into the second equation gives $y = 8$, and substituting $z = 2$ and $y = 8$ into the first equation yields $x = -15$. Therefore, the solution set is the single point $(x, y, z) = (-15, 8, 2)$.

Gaussian elimination

Elimination of variables method is natural for solving by hand, but is a little cumbersome to implement it as a computer program. The method of Gaussian elimination can be described as an easy to implement algorithm and is the standard algorithm for numerically solving a system of linear equations. Gaussian elimination relies on row-reduction of an augmented matrix as explained below.

Augmented matrix

Given a system of linear equations represented in the matrix form as $\mathbf{Ax} = \mathbf{b}$, its augmented matrix is obtained by concatenating m-by-1 column vector $\mathbf{b}$ to m-by-n $\mathbf{A}$ and creating m-by-$(n+1)$ augmented matrix $[\mathbf{A} \mid \mathbf{b}]$.

For example,

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & -1 \\ 2 & -2 & 4 \\ -1 & 1/2 & -1 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \quad \Rightarrow \quad [\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 3 & 2 & -1 & 1 \\ 2 & -2 & 4 & -2 \\ -1 & 1/2 & -1 & 0 \end{array} \right]$$

Row reduction

In **row reduction**, the linear system is represented as an augmented matrix and this matrix is then modified using elementary row operations until it reaches reduced row echelon form. A matrix is in **row echelon form** if

- All nonzero rows (rows with at least one nonzero element) are above any rows of all zeroes (all zero rows, if any, belong at the bottom of the matrix).
- The leading coefficient (the first nonzero number from the left, also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.

This is an example of a 3×5 matrix in row echelon form:

$$\begin{bmatrix} 1 & a_0 & a_1 & a_2 & a_3 \\ 0 & 0 & 2 & a_4 & a_5 \\ 0 & 0 & 0 & 1 & a_6 \end{bmatrix}$$

A matrix is in **reduced row echelon form** (also called **row canonical form**) if it satisfies the following conditions:

- It is in row echelon form.
- Every leading coefficient is 1 and is the only nonzero entry in its column.

This is an example of a matrix in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 & a_1 & 0 & b_1 \\ 0 & 1 & a_2 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{bmatrix}$$

Because these operations are reversible, the augmented matrix produced always represents a linear system that is equivalent to the original.

The Algorithm

Gaussian elimination relies on the following 3 observations:
(1) multiplying any equation by a scalar does not change the solution,
(2) adding any equation to any other equation does not change the solution,
(3) switching any two equations does not change the solution.
These three operations are equivalent to the elementary row operations described above. The main idea of the Gaussian elimination is to use the elementary row operations to simplify the augmented matrix to its reduced row echelon form. Then, the augmented matrix is further simplified such that its first part is an identity matrix. Given this simplified form, the solution is the last column of the augmented matrix. Let us look at examples:

$$\begin{aligned} 2x + y - z &= 8 \quad &(L_1) \\ -3x - y + 2z &= -11 \quad &(L_2) \\ -2x + y + 2z &= -3 \quad &(L_3) \end{aligned}$$

The table below is the row reduction process applied simultaneously to the system of equations, and its associated augmented matrix.
In practice, one does not usually deal with the systems in terms of equations but instead makes use of the augmented matrix, which is more suitable for computer manipulations. The row reduction procedure in the table may be summarized as follows:
eliminate x from all equations below L1, and then eliminate y from all equations below L 2. This will put the system into triangular form. Then, using back-substitution, each unknown can be solved for.

System of equations	Row operations	Augmented matrix
$2x + y - z = 8$ $-3x - y + 2z = -11$ $-2x + y + 2z = -3$		$\begin{bmatrix} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{bmatrix}$
$2x + y - z = 8$ $\dfrac{1}{2}y + \dfrac{1}{2}z = 1$ $2y + z = 5$	$L_2 + \dfrac{3}{2}L_1 \rightarrow L_2$ $L_3 + L_1 \rightarrow L_3$	$\begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 1/2 & 1/2 & 1 \\ 0 & 2 & 1 & 5 \end{bmatrix}$
$2x + y - z = 8$ $\dfrac{1}{2}y + \dfrac{1}{2}z = 1$ $-z = 1$	$L_3 + -4L_2 \rightarrow L_3$	$\begin{bmatrix} 2 & 1 & -1 & 8 \\ 0 & 1/2 & 1/2 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix}$
The matrix is now in echelon form (also called triangular form)		
$2x + y = 7$ $\dfrac{1}{2}y = \dfrac{3}{2}$ $-z = 1$	$L_2 + \dfrac{1}{2}L_3 \rightarrow L_2$ $L_1 - L_3 \rightarrow L_1$	$\begin{bmatrix} 2 & 1 & 0 & 7 \\ 0 & 1/2 & 0 & 3/2 \\ 0 & 0 & -1 & 1 \end{bmatrix}$
$2x + y = 7$ $y = 3$ $z = -1$	$2L_2 \rightarrow L_2$ $-L_3 \rightarrow L_3$	$\begin{bmatrix} 2 & 1 & 0 & 7 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$
$x = 2$ $y = 3$ $z = -1$	$L_1 - L_2 \rightarrow L_1$ $\dfrac{1}{2}L_1 \rightarrow L_1$	$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$

Once *y* is also eliminated from the third row, the result is a system of linear equations in triangular form, and so the first part of the algorithm is complete. From a computational point of view, it is faster to solve the variables in reverse order, a process known as back-substitution. One sees the solution is $z = -1$, $y = 3$, and $x = 2$. So there is a unique solution to the original system of equations.

Instead of stopping once the matrix is in echelon form, one could continue until the matrix is in *reduced* **row echelon form**, as it is done in the table. The process of row reducing until the matrix is reduced is sometimes referred to as **Gauss-Jordan elimination**, to distinguish it from stopping after reaching echelon form.

Another example is for the system we solved before by elimination of variables:

$$
\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 3 & 5 & 6 & | & 7 \\ 2 & 4 & 3 & | & 8 \end{bmatrix} \sim
\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & -4 & 12 & | & -8 \\ 2 & 4 & 3 & | & 8 \end{bmatrix} \sim
\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & -4 & 12 & | & -8 \\ 0 & -2 & 7 & | & -2 \end{bmatrix} \sim
\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & -3 & | & 2 \\ 0 & -2 & 7 & | & -2 \end{bmatrix}
$$

$$
\sim \begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & -3 & | & 2 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \sim
\begin{bmatrix} 1 & 3 & -2 & | & 5 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \sim
\begin{bmatrix} 1 & 3 & 0 & | & 9 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix} \sim
\begin{bmatrix} 1 & 0 & 0 & | & -15 \\ 0 & 1 & 0 & | & 8 \\ 0 & 0 & 1 & | & 2 \end{bmatrix}.
$$

The last matrix is in reduced row echelon form, and represents the solution $x = -15$, $y = 8$, $z = 2$.

Time complexity

The number of arithmetic operations required to perform row reduction is one way of measuring the algorithm's computational efficiency. For example, to solve a system of n equations for n unknowns by performing row operations on the matrix until it is in echelon form, and then solving for each unknown in reverse order, requires $n(n+1)/2$ divisions, $(2n^3 + 3n^2 - 5n)/6$ multiplications, and $(2n^3 + 3n^2 - 5n)/6$ subtractions, for a total of approximately $2n^3/3$ operations. Thus it has arithmetic complexity of $O(n^3)$.

Matrix solution

We are jumping ahead here, this paragraph will be clear after learning about the inverse matrix. If the equation system is expressed in the matrix form $\mathbf{Ax} = \mathbf{b}$, the entire solution set can also be expressed in matrix form. If the matrix $\mathbf{A}$ is square (has m rows and $n=m$ columns) and has full rank (all m rows are independent), then the system has a unique solution given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ where $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$.

4. Rank of a matrix

In <u>linear algebra</u>, the **rank** of a <u>matrix</u> A is the <mark>size of the largest collection of <u>linearly independent</u> columns of</mark> **A** (the **column rank**) <mark>or the size of the largest collection of linearly independent rows of</mark> **A** (the **row rank**).
For every matrix, <mark>the column rank is equal to the row rank</mark>. The rank is commonly denoted $rank(\mathbf{A})$.

The vectors $v_1, v_2, ..., v_k$ are said to be <mark>linearly dependent</mark> if there exist a finite number of scalars $a_1, a_2, ..., a_k$, not all zero, such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0,$$

where zero denotes the zero vector. Otherwise, we said that vectors $v_1, v_2, ..., v_k$ are <mark>linearly independent</mark>.

In particular, 3 vectors u, v, w are linearly independent if $au + bv + cw = 0$ has only one solution $a = b = c = 0$, which is equivalent to the fact that any of the vectors u, v, w is outside of the span of the other two, i.e., $w \neq au + bv$ for any combination of scalars a, b.
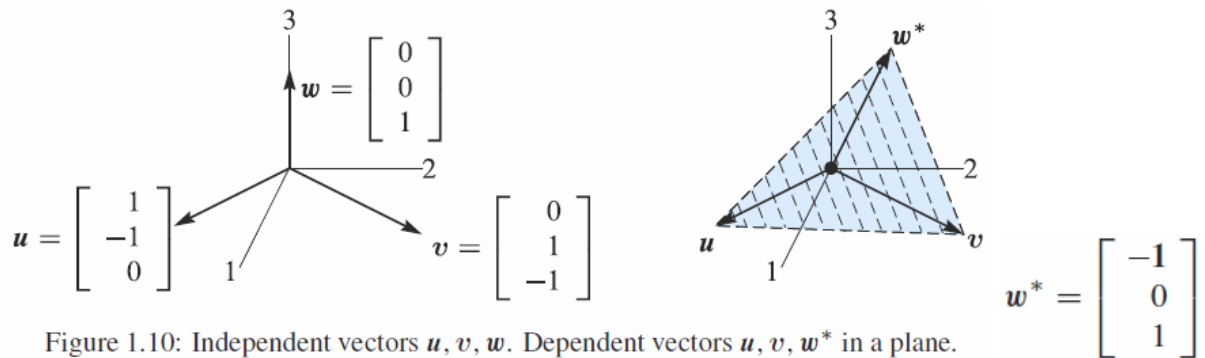

Figure 1.10: Independent vectors u, v, w. Dependent vectors u, v, w^* in a plane.
Independent columns: $Ax = 0$ has one solution. A is an <mark>**invertible matrix.**</mark>
Dependent columns: $Ax = 0$ has many solutions. A is a <mark>**singular matrix.**</mark>

The Gaussian elimination algorithm can be applied to any m-by-n matrix **A**. In this way, for example, some matrices can be transformed to a matrix that has a row echelon form like

$$T = \begin{bmatrix} a & * & * & * & * & * & * & * & * \\ 0 & 0 & b & * & * & * & * & * & * \\ 0 & 0 & 0 & c & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & d & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where the *s are arbitrary entries and a, b, c, d, e are nonzero entries. This echelon matrix **T** contains a wealth of information about **A**: the <u>rank</u> of **A** is 5 since there are 5 non-zero rows in **T**.

5. Linear transformations

Matrices and matrix multiplication reveal their essential features when related to *linear transformations*, also known as *linear maps*. *A real m-by-n matrix* **A** *gives rise to a* <mark>*linear transformation* $R^n \to R^m$ *mapping each vector* **x** *in* R^n *to the (matrix) product* **Ax**, *which is a vector in* R^m.</mark>

Conversely, <mark>*each linear transformation f:* $R^n \to R^m$ *arises from a unique m-by-n matrix* **A**:</mark> *explicitly, the (i, j)-entry of* **A** *is the* i^{th} *coordinate of* $f(e_j)$, *where* $e_j = (0,...,0,1,0,...,0)$ *is the* <u>unit vector</u> *with 1 in the* j^{th} *position and 0 elsewhere.* The matrix **A** *is said to represent the linear map f, and* **A** *is called the transformation matrix of f.*

For example, the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

can be viewed as the transform of the <u>unit square</u> into a <u>parallelogram</u> with vertices at $(0, 0)$, (a, b), $(a + c, b + d)$, and (c, d).

The parallelogram pictured below is obtained by multiplying **A** with each of the column vectors $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
These vectors define the vertices of the unit square.
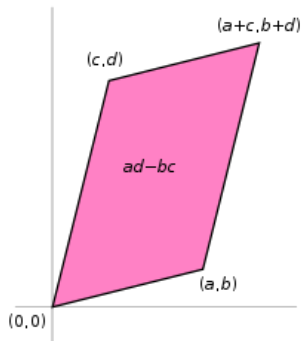
The vectors represented by a 2-by-2 matrix correspond to the sides of a unit square transformed into a parallelogram.

The following table shows a number of 2-by-2 matrices with the associated linear maps of $\mathbf{R}^2$. The blue original is mapped to the green grid and shapes. The origin (0,0) is marked with a black point.

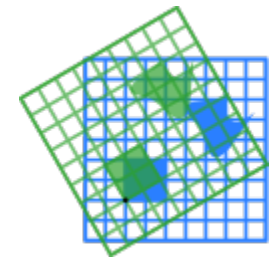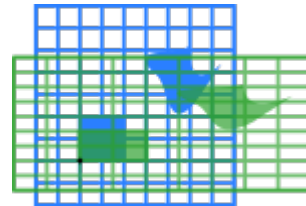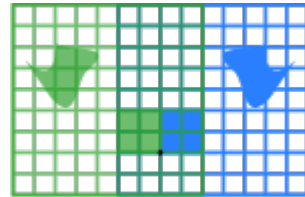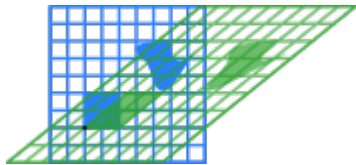| Horizontal shear with m=1.25.	Horizontal flip	Squeeze mapping with r=3/2	Scaling by a factor of 3/2	Rotation by $\pi/6^R = 30°$
$\begin{bmatrix} 1 & 1.25 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 3/2 & 0 \\ 0 & 2/3 \end{bmatrix}$	$\begin{bmatrix} 3/2 & 0 \\ 0 & 3/2 \end{bmatrix}$	$\begin{bmatrix} \cos(\pi/6^R) & -\sin(\pi/6^R) \\ \sin(\pi/6^R) & \cos(\pi/6^R) \end{bmatrix}$

6. Square matrices

A square matrix is a matrix with the same number of rows and columns. An *n-by-n* matrix is known as a square matrix of order *n*. Any two square matrices of the same order can be added and multiplied. The entries a_{ii} form the main diagonal of a square matrix. They lie on the imaginary line which runs from the top left corner to the bottom right corner of the matrix.

Main types

Name	Example with $n = 3$
Diagonal matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$
Lower triangular matrix	$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$
Upper triangular matrix	$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$

The identity matrix $\mathbf{I}_n$ of size *n* is the *n-by-n* matrix in which all elements on the main diagonal are equal to 1 and all other elements are equal to 0:

$$I_1 = \begin{bmatrix} 1 \end{bmatrix}, \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \cdots, \quad I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

It is called identity matrix because multiplication with it leaves a matrix unchanged:

$\mathbf{AI}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A}$ for any *m-by-n* matrix $\mathbf{A}$.

A square matrix $\mathbf{A}$ that is equal to its transpose, i.e., $\mathbf{A} = \mathbf{A}^T$, is a symmetric matrix.

6.1 Invertible matrix and its inverse

A square matrix **A** is called *invertible* or *non-singular* if there exists a matrix **B** such that

$$AB = BA = I_n.$$

If **B** exists, it is unique and is called the *inverse matrix* of **A**, denoted A^{-1}.

A square nxn matrix is **invertible** iff its **determinant** (defined below) is not zero.
A square nxn matrix is **invertible** iff its **rank** is n.
A square nxn matrix A is **invertible** iff the **equation** Ax=0 has only the trivial solution x=0.

If A and B are invertible, then so is A*B: $(A*B)^{-1} = B^{-1} * A^{-1}$.

Finding the inverse of a matrix

A 2 by 2 matrix is invertible if and only if *ad - bc* is not zero:

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

A diagonal matrix has an inverse provided no diagonal entries are zero:

$$A = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \quad \text{then} \quad A^{-1} = \begin{bmatrix} 1/d_1 & & \\ & \ddots & \\ & & 1/d_n \end{bmatrix}$$

A variant of Gaussian elimination called Gauss–Jordan elimination can be used for finding the inverse of a matrix, if it exists.
If **A** is a n by n square matrix, then one can use row reduction to compute its inverse matrix, if it exists.
First, the n by n identity matrix is augmented to the right of **A**, forming a n by $2n$ block matrix [**A** | **I**].
This is because we are trying to find matrix **X** which is the solution of A*X= **I**.
Now through application of elementary row operations, find the reduced echelon form of this n by $2n$ matrix.
The matrix **A** is invertible if and only if the left block can be reduced to the identity matrix **I**; in this case the right block of the final matrix is A^{-1}.
If the algorithm is unable to reduce the left block to **I**, then **A** is not invertible.

For example, consider the following matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

To find the inverse of this matrix, one takes the following matrix augmented by the identity, and row reduces it as a 3 by 6 matrix:

$$[A|I] = \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array} \right].$$

By performing row operations, one can check that the reduced row echelon form of this augmented matrix is:

$$[I|B] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array} \right].$$

The matrix on the left is the identity, which shows A is invertible. The 3 by 3 matrix on the right, B, is the inverse of A. This procedure for finding the inverse works for square matrices of any size.

6.2. Trace

The trace, tr($\mathbf{A}$) of a square matrix $\mathbf{A}$ is the sum of its diagonal entries, tr($\mathbf{A}$) = sum(A_{ii}). While matrix multiplication is not commutative as mentioned above, the trace of the product of two matrices is independent of the order of the factors:

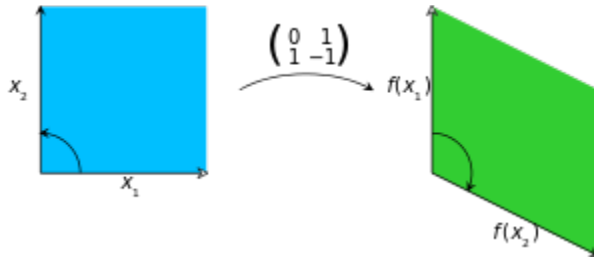tr($\mathbf{AB}$) = tr($\mathbf{BA}$).

This is immediate from the definition of matrix multiplication:

$$\text{tr}(\mathbf{AB}) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ji} = \text{tr}(\mathbf{BA}).$$

Also, the trace of a matrix is equal to that of its transpose, i.e.,

tr($\mathbf{A}$) = tr($\mathbf{A}^{\mathrm{T}}$).

6.3. Determinant

A linear transformation on $\mathbf{R}^2$ is given by the indicated matrix. The determinant of this matrix is -1, as the area of the green parallelogram at the right is 1, but the map reverses the orientation, since it turns the counterclockwise orientation of the vectors to a clockwise one.

The *determinant det*($\mathbf{A}$) or $|\mathbf{A}|$ of a square matrix $\mathbf{A}$ is a number encoding certain properties of the matrix.

A matrix is invertible if and only if its determinant is nonzero.

Its absolute value equals the area (in $\mathbf{R}^2$) or volume (in $\mathbf{R}^3$) of the image of the unit square (or cube),
while its sign corresponds to the orientation of the corresponding linear map: the determinant is positive if and only if the orientation is preserved.

The determinant of 2-by-2 matrices is given by

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

The determinant of 3-by-3 matrices involves 6 terms (rule of Sarrus):

Sarrus' rule: The determinant of the three columns on the left is the sum of the products along the solid diagonals minus the sum of the products along the dashed diagonals

Definition of Determinant

There are various ways to define the determinant of a square matrix **A**. Perhaps the most natural way is expressed in terms of the columns of the matrix. If we write an $n \times n$ matrix in terms of its column vectors

$$A = \begin{bmatrix} a_1, & a_2, & \ldots, & a_n \end{bmatrix}$$

where the a_j are vectors of size n, then the determinant of **A** is defined so that

$$\det \begin{bmatrix} a_1, & \ldots, & ba_j + cv, & \ldots, a_n \end{bmatrix} = b \det(A) + c \det \begin{bmatrix} a_1, & \ldots, & v, & \ldots, a_n \end{bmatrix}$$

$$\det \begin{bmatrix} a_1, & \ldots, & a_j, & a_{j+1}, & \ldots, a_n \end{bmatrix} = - \det \begin{bmatrix} a_1, & \ldots, & a_{j+1}, & a_j, & \ldots, a_n \end{bmatrix}$$

$$\det(I) = 1$$

where b and c are scalars, v is any vector of size n and I is the identity matrix of size n. Adding a multiple of any row to another row, or a multiple of any column to another column, does not change the determinant. Interchanging two rows or two columns affects the determinant by multiplying it by -1. Using these operations, any matrix can be transformed to a lower (or upper) triangular matrix, and for such matrices the determinant equals the product of the entries on the main diagonal; this provides a method to calculate the determinant of any matrix.

An idea of Gaussian elimination can be used to find determinant of a matrix. For example, the determinant of

$$A = \begin{bmatrix} -2 & 2 & -3 \\ -1 & 1 & 3 \\ 2 & 0 & -1 \end{bmatrix}$$

can be computed using the following matrices:

$$B = \begin{bmatrix} -2 & 2 & -3 \\ 0 & 0 & 4.5 \\ 2 & 0 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} -2 & 2 & -3 \\ 0 & 0 & 4.5 \\ 0 & 2 & -4 \end{bmatrix}, \quad D = \begin{bmatrix} -2 & 2 & -3 \\ 0 & 2 & -4 \\ 0 & 0 & 4.5 \end{bmatrix}.$$

Site http://www.math.odu.edu/~bogacki/cgi-bin/lat.cgi can be used to see step by step calculations for matrices that are based on the idea of Gaussian elimination. It shows how reduction to row echelon form works and how it helps in

- Solving a system of linear equations
- Finding an inverse matrix
- Finding a determinant
- Finding a rank

Properties of the determinant

The determinant has many properties. Some basic properties of determinants are:

1. $\det(I_n) = 1$ where *In* is the $n \times n$ identity matrix.
2. $\det(A^T) = \det(A).$
3. $\det(A^{-1}) = \dfrac{1}{\det(A)} = \det(A)^{-1}.$
4. For square matrices A and B of equal size,

$$\det(AB) = \det(A)\det(B).$$

5. $\det(cA) = c^n \det(A)$ for an $n \times n$ matrix.
6. If A is a triangular matrix, i.e. $a_{i,j} = 0$ whenever $i > j$ or, alternatively, whenever $i < j$, then its determinant equals the product of the diagonal entries:

$$\det(A) = a_{1,1}a_{2,2}\cdots a_{n,n} = \prod_{i=1}^{n} a_{i,i}.$$

A number of additional properties relate to the effects on the determinant of changing particular rows or columns:

7. If in a matrix, any row or column is 0, then the determinant of that particular matrix is 0.
8. Whenever two columns of a matrix are identical, or more generally some column can be expressed as a linear combination of the other columns (i.e. the columns of the matrix form a linearly dependent set), its determinant is 0.
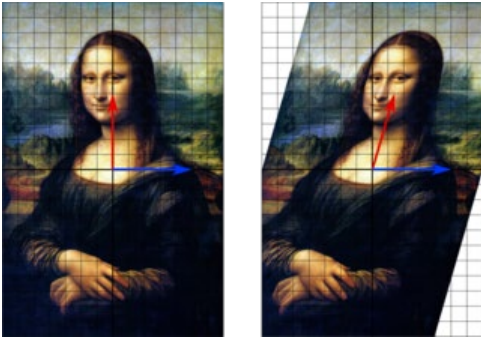
6.4. Eigenvalues and eigenvectors

In linear algebra, an **eigenvector** or **characteristic vector** of a square matrix is a vector that points in a direction which is invariant under the associated linear transformation. In other words, if $\mathbf{v}$ is a vector which is not zero, then it is an eigenvector of a square matrix A if $A\mathbf{v}$ is a scalar multiple of $\mathbf{v}$. This condition could be written as the equation

$$A\mathbf{v} = \lambda\mathbf{v},$$

where λ is a number (also called a scalar) known as the **eigenvalue** or **characteristic value** associated with the eigenvector $\mathbf{v}$.

There is a correspondence between n by n square matrices and linear transformation from an n-dimensional vector space to itself. For this reason, it is equivalent to define eigenvalues and eigenvectors using either the language of matrices or the language of linear transformations.

In this shear mapping the red arrow changes direction but the blue arrow does not. The blue arrow is an eigenvector of this shear mapping because it doesn't change direction, and since in this example its length is unchanged its eigenvalue is 1.

Two dimensional example

Consider the transformation matrix $\mathbf{A}$, given by,

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

The eigenvectors $\mathbf{v}$ of this transformation satisfy the equation,

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Rearrange this equation to obtain

$$(A - \lambda I)\mathbf{v} = 0,$$

which has a nontrivial solution ($\mathbf{v} \neq 0$) only when its determinant $|A - \lambda I|$ equals zero.

Set the determinant to zero to obtain the polynomial equation,

$$p(\lambda) = |A - \lambda I| = 3 - 4\lambda + \lambda^2 = 0,$$

known as the characteristic polynomial of the matrix $\mathbf{A}$. In this case, it has the roots $\lambda = 1$ and $\lambda = 3$.

For $\lambda = 1$, the equation becomes,

$$(A - I)\mathbf{v} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{Bmatrix} v_1 \\ v_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix},$$

which has the solution,

$$\mathbf{v} = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}.$$

For $\lambda = 3$, the equation becomes,

$$(A - 3I)\mathbf{w} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{Bmatrix} w_1 \\ w_2 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix},$$

which has the solution,

$$\mathbf{w} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}.$$

Thus, the vectors $\mathbf{v}$ and $\mathbf{w}$ are eigenvectors of $\mathbf{A}$ associated with the eigenvalues $\lambda = 1$ and $\lambda = 3$, respectively.

Diagonal matrices

Matrices with entries only along the main diagonal are called *diagonal matrices*. It is easy to see that the eigenvalues of a diagonal matrix are the diagonal elements themselves. Consider the matrix **A**,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

The characteristic polynomial of A is given by

$$p(\lambda) = |A - \lambda I| = (1 - \lambda)(2 - \lambda)(3 - \lambda) = 0,$$

which has the roots $\lambda = 1$, $\lambda = 2$ and $\lambda = 3$.

Associated with these roots are the eigenvectors,

$$\mathbf{u} = \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix}, \quad \mathbf{v} = \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix}, \quad \text{and} \quad \mathbf{w} = \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix},$$

respectively.

Triangular matrices

Consider the lower triangular matrix A,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 3 \end{bmatrix}.$$

The characteristic polynomial of A is given by

$$p(\lambda) = |A - \lambda I| = (1 - \lambda)(2 - \lambda)(3 - \lambda) = 0,$$

which has the roots $\lambda = 1$, $\lambda = 2$ and $\lambda = 3$.

Associated with these roots are the eigenvectors,

$$\mathbf{u} = \begin{Bmatrix} 1 \\ -1 \\ 1/2 \end{Bmatrix}, \quad \mathbf{v} = \begin{Bmatrix} 0 \\ 1 \\ -3 \end{Bmatrix}, \quad \text{and} \quad \mathbf{w} = \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix},$$

respectively.

Eigenvalues

Since the eigenvalues are roots of the characteristic polynomial, an $n \times n$ matrix has at most n distinct eigenvalues. If the matrix has real entries, the coefficients of the characteristic polynomial are all real; but it may have fewer than n real roots, or no real roots at all.

For example, consider the [cyclic permutation matrix](#)

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

This matrix shifts the coordinates of the vector up by one position, and moves the first coordinate to the bottom. Its characteristic polynomial is $1 - \lambda^3$ which has only one real root $\lambda_1 = 1$. Any vector with three equal non-zero coordinates is an eigenvector for this eigenvalue. For example,

$$A \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix} = 1 \cdot \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$$

Properties of eigenvalues and eigenvectors

Let **A** be an arbitrary n-by-n matrix with eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_n$. (Here it is understood that an eigenvalue with algebraic multiplicity μ occurs μ times in this list.) Then

- The [trace](#) of A, defined as the sum of its diagonal elements, is also the sum of all eigenvalues:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$
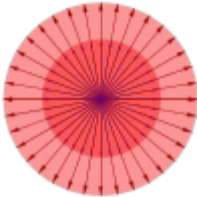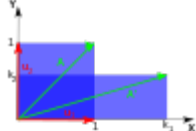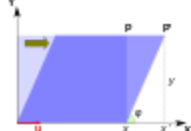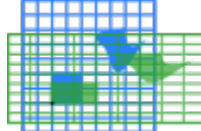
- The [determinant](#) of A is the product of all eigenvalues:

$$\det(A) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \lambda_2 \cdots \lambda_n$$

- The eigenvalues of the kth power of A, i.e. the eigenvalues of A^k, for any positive integer k, are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$
- The matrix A is invertible if and only if all the eigenvalues λ_i are nonzero.
- If A is invertible, then the eigenvalues of A^{-1} are $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n$. Clearly, the geometric multiplicities coincide. Moreover, since the characteristic polynomial of the inverse is the [reciprocal polynomial](#) for that of the original, they share the same algebraic multiplicity.
- If A is a [symmetric](#) real matrix then every eigenvalue is real

Eigenvalues of geometric transformations

The following table presents some example transformations in the plane along with their 2×2 matrices, eigenvalues, and eigenvectors.

	scaling	unequal scaling	rotation	horizontal shear	hyperbolic rotation
illustration					
matrix	$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$	$\begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$	$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ $c = \cos\theta$ $s = \sin\theta$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} c & s \\ s & c \end{bmatrix}$ $c = \cosh\varphi$ $s = \sinh\varphi$
characteristic polynomial	$(\lambda - k)^2$	$(\lambda - k_1)(\lambda - k_2)$	$\lambda^2 - 2c\lambda + 1$	$(\lambda - 1)^2$	$\lambda^2 - 2c\lambda + 1$
eigenvalues λ_i	$\lambda_1 = \lambda_2 = k$	$\lambda_1 = k_1$ $\lambda_2 = k_2$	$\lambda_1 = e^{i\theta} = c + si$ $\lambda_2 = e^{-i\theta} = c - si$	$\lambda_1 = \lambda_2 = 1$	$\lambda_1 = e^{\varphi}$ $\lambda_2 = e^{-\varphi}$,
eigenvectors	All non-zero vectors	$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$ $u_2 = \begin{bmatrix} 1 \\ +i \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $u_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Diagonalization and eigendecomposition

Let A have a set of n linearly independent eigenvectors. Let Q be a square matrix whose columns are those eigenvectors, in any order. Then we will have $AQ = Q\Lambda$, where Λ is the diagonal matrix such that Λ_{ii} is the eigenvalue associated to column i of Q. Since the columns of Q are linearly independent, the matrix Q is invertible. Premultiplying both sides by Q^{-1} we get $Q^{-1}AQ = \Lambda$. By definition, therefore, the matrix A is diagonalizable.

If A is diagonalizable, the space of all n-coordinate vectors can be decomposed into the direct sum of the eigenspaces of A. This decomposition is called the eigendecomposition of A, and it is preserved under change of coordinates.

Finite-dimensional spectral theorem. Any symmetric matrix whose entries are real can be diagonalized by an orthogonal matrix. More explicitly: For every symmetric real matrix $\mathbf{A}$ there exists a real orthogonal matrix $\mathbf{Q}$ such that $\mathbf{D} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}$ is a diagonal matrix. Every symmetric matrix is thus, up to choice of an orthonormal basis, a diagonal matrix. $\mathbf{D}$ is composed of the eigenvalues of $\mathbf{A}$.

Orthogonal matrix

An *orthogonal matrix* is a square matrix with real entries whose columns and rows are orthogonal unit vectors (i.e., orthonormal vectors). Equivalently, a matrix A is orthogonal if its transpose is equal to its inverse:

$$A^{\mathrm{T}} = A^{-1},$$

which entails

$$A^{\mathrm{T}}A = AA^{\mathrm{T}} = I,$$

where I is the identity matrix.

An orthogonal matrix A is necessarily invertible (with inverse $A^{-1} = A^{\mathrm{T}}$). The determinant of any orthogonal matrix is either $+1$ or -1. A *special orthogonal matrix* is an orthogonal matrix with determinant $+1$. As a linear transformation, every orthogonal matrix with determinant $+1$ is a pure rotation, while every orthogonal matrix with determinant -1 is either a pure reflection, or a composition of reflection and rotation.

Calculation of eigenvalues and eigenvectors

The eigenvalues of a matrix **A** can be determined by finding the roots of the characteristic polynomial. Explicit algebraic formulas for the roots of a polynomial exist only if the degree n is 4 or less. Therefore, for matrices of order 5 or more, the eigenvalues and eigenvectors cannot be obtained by an explicit algebraic formula, and must therefore be computed by approximate numerical methods.

In theory, the coefficients of the characteristic polynomial can be computed exactly, since they are sums of products of matrix elements; and there are algorithms that can find all the roots of a polynomial of arbitrary degree to any required accuracy. However, this approach is not viable in practice because the coefficients would be contaminated by unavoidable round-off errors, and the roots of a polynomial can be an extremely sensitive function of the coefficients. Efficient, accurate methods to compute eigenvalues and eigenvectors of arbitrary matrices were not known until the advent of the QR algorithm in 1961.

Once the (exact) value of an eigenvalue is known, the corresponding eigenvectors can be found by finding non-zero solutions of the eigenvalue equation, that becomes a system of linear equations with known coefficients.

6.5. Few other types of square matrices

Definite matrix

A symmetric $n{\times}n$-matrix is called *positive-definite* (respectively negative-definite; indefinite), if for all nonzero vectors $\mathbf{x} \in \mathbf{R}^n$ the associated quadratic form given by

$$Q(x) = x^T A x$$

takes only positive values (respectively only negative values; both some negative and some positive values). If the quadratic form takes only non-negative (respectively only non-positive) values, the symmetric matrix is called positive-semidefinite (respectively negative-semidefinite); hence the matrix is indefinite precisely when it is neither positive-semidefinite nor negative-semidefinite.

- If A is also positive-definite, positive-semidefinite, negative-definite, or negative-semidefinite every eigenvalue is positive, non-negative, negative, or non-positive respectively.