Reminder. A matrix $A \in \mathbb{R}^{n \times n}$ is called invertible, if there exist a matrix $B \in \mathbb{R}^{n \times n}$ such that $BA = \text{Id}_n$ and $AB = \text{Id}_n$. $B$ is called the inverse (matrix) of $A$, and we write $A^{-1}$ for it.

**Theorem 8.1. Properties of matrix inverses:**

a) If $B = A^{-1}$, then $B$ is invertible and $B^{-1} = A$, i.e. $(A^{-1})^{-1} = A$

b) If $A$ and $B$ are invertible, and $AB$ exists, then $AB$ is invertible, and $(AB)^{-1} = B^{-1}A^{-1} \quad \leftarrow \text{note: reverse order!!}

c) If $A$ is invertible, then $A^\top$ is invertible, and $(A^\top)^{-1} = (A^{-1})^\top$

Proof: exercise

**What is that good for?**

**Remark (Solving systems of linear equations).** Assume $AX = B$ and we know $A$ and $B$, but not $X$. If $A$ is invertible,

- a system is solvable for arbitrary $B$ (of the right size),
- the solution is uniquely determined, and
- we can write down the solution in one line:

\[
X = \text{Id} X = A^{-1}AX = A^{-1}B.
\]

Note: $Ax = b$ (solving a system of linear equations) is a special case: $b = A^{-1}x$
Example 8.2. tRNA (from Otto&Day, 2007)
tRNA is used in the process of translating mRNA into proteins. tRNA comes in two states: unbound, or bound to an amino acid. We call the number of unbound molecules $n_1$ and the number of bound molecules $n_2$. Their dynamics can be modelled with four parameters $\alpha, \beta, \gamma, \delta \geq 0$:

- $\alpha$: the number of molecules produced per time unit (by transcription)
- $\beta$: the binding rate (i.e. unbound tRNA turning into bound tRNA)
- $\gamma$: the rate of amino acid loss (i.e. bound tRNA becomes unbound)
- $\delta$: the rate of tRNA degradation (by making a protein)

as

\[
\begin{aligned}
  n_{1}^{\text{next}} &= n_{1}^{\text{cur}} + \alpha - \beta n_{1}^{\text{cur}} + \gamma n_{2}^{\text{cur}} - \delta n_{1}^{\text{cur}} \\
  n_{2}^{\text{next}} &= n_{2}^{\text{cur}} + \beta n_{1}^{\text{cur}} - \gamma n_{2}^{\text{cur}} - \delta n_{2}^{\text{cur}}
\end{aligned}
\]

we can write this with vectors and matrices

\[
\begin{pmatrix}
  n_{1}^{\text{next}} \\
  n_{2}^{\text{next}}
\end{pmatrix}
=:
\begin{pmatrix}
  n_{1}^{\text{cur}} \\
  n_{2}^{\text{cur}}
\end{pmatrix} +
\begin{pmatrix}
  -\beta & \gamma \\
  \beta & -\gamma - \delta
\end{pmatrix}
\begin{pmatrix}
  n_{1}^{\text{cur}} \\
  n_{2}^{\text{cur}}
\end{pmatrix} -
\begin{pmatrix}
  -\alpha \\
  0
\end{pmatrix}
=:
\begin{pmatrix}
  \tilde{n} \\
  \nu
\end{pmatrix}
\]

so

\[
\tilde{n} - n = Mn - \nu.
\]

In a stable state, we’ll have $\tilde{n} = n$, so

\[
Mn = \nu.
\]

If $M$ is invertible, we can use its inverse to solve for the stable state $n$:

\[
n = M^{-1}\nu
\]

But is $M$ really invertible? We have to check $(-\beta - \delta)(-\gamma - \delta) - \gamma\beta \neq 0$.

\[
(-\beta - \delta)(-\gamma - \delta) - \gamma\beta = \beta\gamma + \delta\gamma + \beta\delta + \delta^2 - \beta\gamma
\]

\[
= \delta\gamma + \beta\delta + \delta^2
\]

this is larger than 0, if $\delta > 0$. This makes sense: if no tRNA is ever degraded, one cannot expect a stable state (except if also $\alpha = 0$, in which case nothing is going on at all).
Observation. Linear Algebra courses typically spend a lot of time on 1) how to check if a matrix is invertible, and – related to this – how to compute the inverse, 2) solving systems of linear equations. We don’t. $A^{-1}$ is a handy symbol to have (for example to write $x = A^{-1}b$ for the solution of a system of linear equations), but in practice (especially numerically), one hardly ever needs it. Note: computers have dedicated routines to solve $AX = B$ for $X$, and these are more efficient and more stable than first computing $A^{-1}$, and then multiplying it with $B$.

- Matlab: $X = \text{linsolve}(A,B)$; or shorthand $X = A\backslash B$;
- Python: $X = \text{solve}(A,B)$ (from numpy.linalg package)
- Matematica: $X = \text{LinearSolve}[A,B]$ 

If you ever do need $A^{-1}$: for $2 \times 2$ or diagonal matrices: use the explicit formula. Beyond that: use a computer (if you know that $A^{-1}$ exists)

- Matlab: $B = \text{inv}(A)$;
- Python: $B = \text{inv}(A)$ (from numpy.linalg package)
- Matematica: $B = \text{Inverse}[A]$

8.1 Inner Product and Norm

Definition 8.3. Inner product Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ be two vectors. Then we call $x^\top y$ (which is a number) the inner product of $x$ and $y$, and we write $\langle x, y \rangle$ for it.

Remark.

- $x^\top y$ is just a number, because a $1 \times n$ matrix multiplied with $n \times 1$ matrix results in a $1 \times 1$ matrix.
- computing the inner product is very easy: $\langle x, y \rangle = x^\top y = \sum_{i=1}^{n} x_i y_i$
- other names for the same object:
  - scalar product, because result is a scalar (a number), not a vector or matrix,
dot product, because another notation for it is \( x \cdot y \) (but we won’t use that)

Lemma 8.4. The inner product is a function \( \langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \) of two vector arguments. It has the following properties:

- symmetry: for all \( x \in \mathbb{R}^n, y \in \mathbb{R}^n \): \( \langle x, y \rangle = \langle y, x \rangle \)
- linearity in the first argument, when the second is held fixed.
  For arbitrary fixed \( z \in \mathbb{R}^n \):
  - \( \forall x, y \in \mathbb{R}^n: \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \), and
  - \( \forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}: \langle \alpha x, z \rangle = \alpha \langle x, z \rangle \).
- linearity in the second argument, when the first is held fixed.
  For arbitrary fixed \( x \in \mathbb{R}^n \):
  - \( \forall y, z \in \mathbb{R}^n: \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \), and
  - \( \forall y \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}: \langle x, \alpha y \rangle = \alpha \langle x, y \rangle \).
- positive definiteness:
  \( \forall x \in \mathbb{R}^n : \langle x, x \rangle \geq 0 \), with \( \langle x, x \rangle = 0 \) if and only if \( x = 0 \).

Proof. Follows directly from the definition: \( \langle x, y \rangle = x^\top y = \sum_{i=1}^{n} x_i y_i \). 

Example 8.5. ISBN. The last digit of an (10-digit) ISBN code (for ordering books) is the check digit: It is computed in the following way:

- let \( x = (x_1, x_2, \ldots, x_9) \) be first the 9 digits of the ISBN code.
- set \( y = (1, 2, 3, 4, 5, 6, 7, 8, 9)^\top \) (always).
- let \( r = \langle x, y \rangle \mod 11 \) (i.e. remainder in integer division by 11).
- the check digit is equal to \( r \), if \( r \in \{0, 1, \ldots, 9\} \), and \( 'X' \) if \( r = 10 \).

Example: somebody tries to order the book with ISBN 1601982682: \( x = (1, 6, 0, 1, 9, 8, 2, 6, 8) \).

\[
y^\top x = 1 + 12 + 0 + 4 + 45 + 48 + 14 + 48 + 72 = 244
\]

\[
r = 244 \mod 11 = 2 \quad \text{(because 244 = 22 \cdot 11 + 2).}
\]

The check digit should be 2. This is the case, the order will be processed.
Example: somebody makes a mistake and orders ISBN 1601892682: \( x = (1, 6, 0, 1, 8, 9, 2, 6, 8) \).

\[
y^\top x = 1 + 12 + 0 + 4 + 40 + 54 + 14 + 48 + 72 = 245
\]

\[
r = 245 \mod 11 = 3 \quad \text{(because } 244 = 22 \cdot 11 + 3).\]

The check digit should be 3, but is 2. An error is detected, order will be declined.

**Definition 8.6.** Vectors \( x, y \in \mathbb{R}^n \) with \( \langle x, y \rangle = 0 \) are called orthogonal (to each other), and we write \( x \perp y \). Alternative name: perpendicular.

**Example 8.7.** In \( \mathbb{R}^2 \) and \( \mathbb{R}^3 \), this definition corresponds to our intuitive understanding of orthogonality:

- \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 0 \\ 1 \end{pmatrix} \)
- \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 0 \\ -1 \end{pmatrix} \)
- \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \perp \begin{pmatrix} 1 \\ -1 \end{pmatrix} \)
- \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \not\perp \begin{pmatrix} 1 \\ 1 \end{pmatrix} \)

**Definition 8.8.** Norm of a vector. Let \( x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \) be a vector. Then \( \sqrt{x^\top x} \) is a non-negative number that we call the norm of \( x \), and we write \( \| x \| \) for it.

**Observation.** The norm of \( x \) is just what in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \) we’re used to calling the length of \( x \):

\[
\| x \| = \sqrt{x^\top x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{8.9}
\]

**Example 8.9.**

- \( \| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \| = \sqrt{1^2 + 0^2} = 1 \),
- \( \| \begin{pmatrix} -1 \\ 0 \end{pmatrix} \| = \sqrt{(-1)^2 + 0^2} = 1 \),
- \( \| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \| = \sqrt{1^2 + 1^2} = \sqrt{2} \),
- \( x = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n, \| x \| = \sqrt{n} \).
Note: often one tries to avoid the $\sqrt{\cdot}$ and uses expression with $\|x\|_2$ instead of $\|x\|$ (because $\|x\|_2^2 = \langle x, x \rangle = x^\top x$)

**Lemma 8.10.** The norm is a function, $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$, of one vector argument. It has the following properties:

- For any $x \in \mathbb{R}$: $\|x\| \geq 0$, with $\|x\| = 0$ if and only if $x = 0$.
- For any $x \in \mathbb{R}, \lambda \in \mathbb{R}$: $\|\lambda x\| = |\lambda|\|x\|$.

**Proof.** 1) $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$. Square roots are never negative, and 0, iff the value inside the square root is 0. $\sum_{i=1}^{n} x_i^2$ is 0, iff all $x_i = 0$.

2) Exercise sheet

**Beware:** $\| \cdot \|$ is NOT LINEAR (see exercise).

**Theorem 8.11.** The inner product of two vectors, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ with $x \neq 0$ and $y \neq 0$ is related to the (geometric) angle between the vectors by the following relation:

$$\langle x, y \rangle = \|x\|\|y\| \cos \angle(x, y) \quad (8.10)$$

**Proof.** We form a triangle with end points 0, $x$ and $y$.

The cosine law of high school, tells us that for any triangle with side length $a, b, c$, and angle $\gamma$ opposite $c$:

$$c^2 = a^2 + b^2 - 2ab \cos \gamma \quad (8.11)$$

(generalization of Pythagorean theorem: if $\gamma = 90^\circ$, then $\cos \gamma = 0$, so $c^2 = a^2 + b^2$).

$$2ab \cos \gamma = a^2 + b^2 - c^2 \quad (8.12)$$

Substituting, $a = \|x\|$, $b = \|y\|$, $c = \|y - x\|$, and $\gamma = \angle(x, y)$, we get

$$2\|x\|\|y\| \cos \angle(x, y) = \|x\|^2 + \|y\|^2 - \|y - x\|^2 \quad (8.13)$$

$$= \langle x, x \rangle + \langle y, y \rangle - \langle y - x, y - x \rangle \quad (8.14)$$
But by the linearity of $\langle \cdot, \cdot \rangle$ in its left and right argument, we have $\langle y - x, y - x \rangle = \langle y, y \rangle - \langle y, x \rangle - \langle x, y \rangle + \langle x, x \rangle$ and symmetry tells us $\langle y, x \rangle = \langle x, y \rangle$,

$$
\langle y - x, y - x \rangle = \langle y, y \rangle - \langle y, x \rangle - \langle x, y \rangle + \langle x, x \rangle = \langle x, x \rangle + \langle y, y \rangle - \langle y, y \rangle + 2 \langle x, y \rangle - \langle x, x \rangle \quad (8.15)
$$

Dividing both sides by 2 yields: $\|x\|\|y\| \cos \angle (x, y) = \langle x, y \rangle$.  

\qed