MAT 460 Numerical Differential Equations

Version 21.10.14

Michael Karow, Andrea Dziubek

Lecture 4

Topics: Matrix Geometrie and

Singular Value Decomposition

A linear map $x \longmapsto Ax$ transforms lines into lines or points.

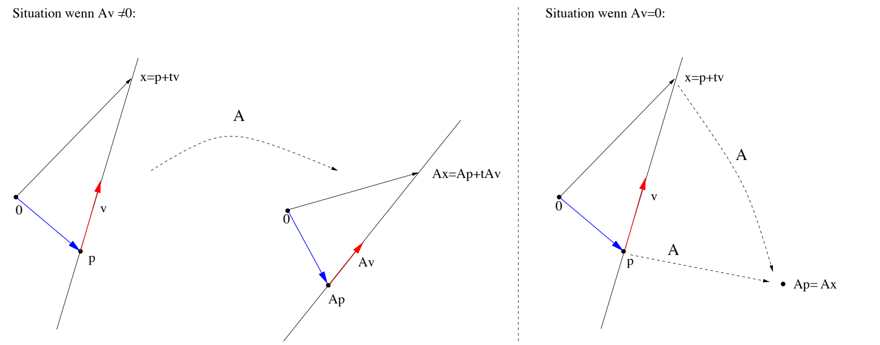
Explanation: The line through the point $p \in \mathbb{R}^n$ in the direction $v \in \mathbb{R}^n, v \neq 0$, is the set of all $x \in \mathbb{R}^n$ with

$$x = p + t\,v, \qquad t \in \mathbb{R}.$$

Multiplication of all x on the line by $A \in \mathbb{R}^{m \times n}$, gives the set of all

$$Ax = Ap + t\,Av, \qquad t \in \mathbb{R}. \qquad (**)$$

This is the line through the point Ap with the direction vector Av, except if $Av = 0$. In this case the image set $(**)$ is only the Ap.



Note: The case $v \neq 0$ and $Av = 0$ is not possible, when A is invertible.

A linear map $x \longmapsto Ax$ transforms parallel lines into parallel lines or points.

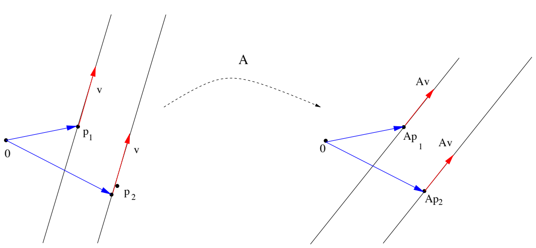
Explanation: Two lines are parallel if they have the same direction vector (or if their direction vectors are multiple of each other), e. g.

$$\text{Line 1:} \qquad x_1 = p_1 + t\,v, \qquad t \in \mathbb{R}.$$
$$\text{Line 2:} \qquad x_2 = p_2 + t\,v, \qquad t \in \mathbb{R}.$$

The image lines are

$$\text{image of line 1:} \qquad Ax_1 = Ap_1 + t\,Av, \qquad t \in \mathbb{R}.$$
$$\text{image of line 2:} \qquad Ax_2 = Ap_2 + t\,Av, \qquad t \in \mathbb{R}.$$

The direction vectors of the image lines are equal. So the image lines are again parallel.

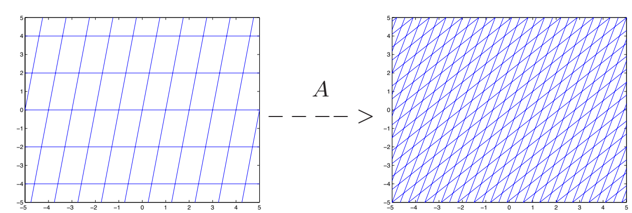


Fact: If a line passes through the origin, then its image also passes through the origin.

What we conclude from these facts?

An invertible linear map $x \mapsto Ax, x \in \mathbb{R}^n$,

transforms a grid of parallel lines into a grid of parallel lines.

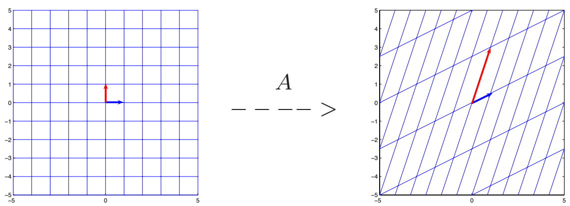


The linear image of the unit grid

The image of the unit grid can be read off from the matrix, i. e.,
The columns of A are the images of the canonical basis vectors.
For the 2×2 case:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{A} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \qquad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix}$$

Illustration:

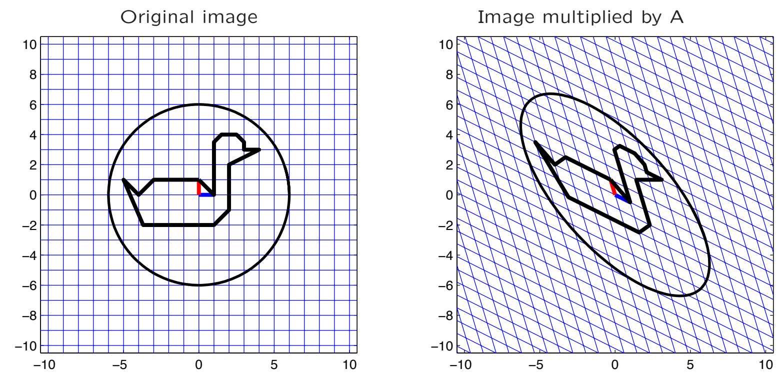


Left Picture: Unit grid with canonical basis vectors.

Right Picture: The A-image of the unit grid with the A-images of the basis vectors.

The matrix in this example is $A = \begin{bmatrix} 1 & 1 \\ 0.5 & 3 \end{bmatrix}$.

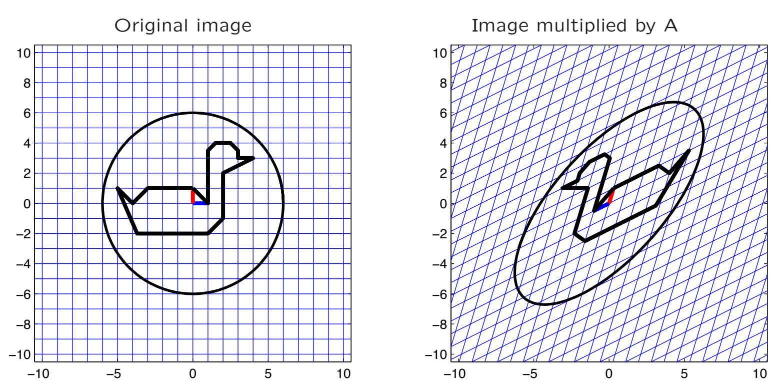
Example: Linear transformation of figures I



The matrix for this example is $A = \begin{bmatrix} 1 & -0.5 \\ -0.3 & 1 \end{bmatrix}$.

Note: The area of an arbitrary figure which is transformed by $x \mapsto Ax$ is multiplied by the factor $|det(A)|$.

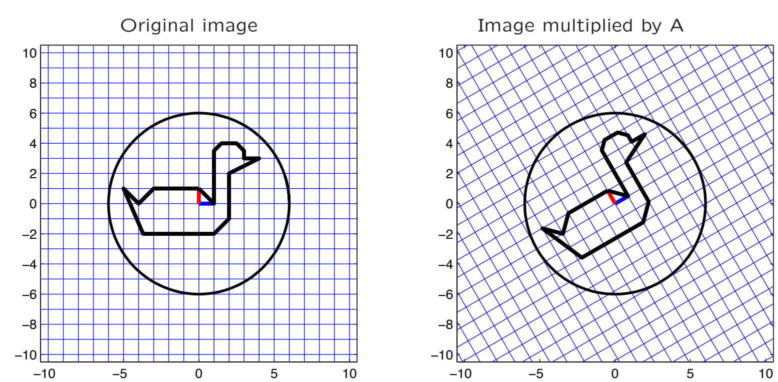
Example: Linear transformations of figures II



The matrix of this example is $A = \begin{bmatrix} -1 & 1 \\ -0.5 & 3 \end{bmatrix}$. We have $det(A) < 0$.

Matrices with negative determinant reverse the orientation.

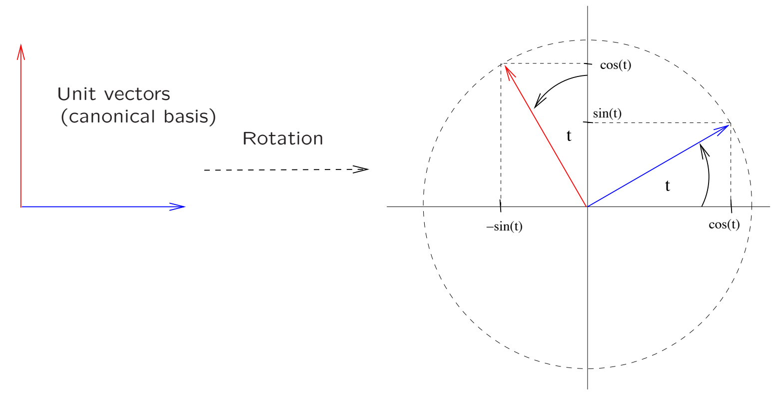
Example: Linear transformations of Figures III



The matrix of this example is $A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$, $\phi = \pi/6$.

A is a **rotation matrix**.

Construction of rotation matrices in $\mathbb{R}^2$



The images of the canonical basis vectors after rotation about the angle t anticlockwise are

$$\begin{bmatrix}\cos(t)\\ \sin(t)\end{bmatrix}, \quad \begin{bmatrix}-\sin(t)\\ \cos(t)\end{bmatrix}.$$

The columns of the corresponding rotation matrix are the images of the basis vectors.

So the rotation matrix is

$$A = \begin{bmatrix}\cos(t) & -\sin(t)\\ \sin(t) & \cos(t)\end{bmatrix}.$$

Rotations in $\mathbb{R}^3$

Let $v \in \mathbb{R}^3$ be a unit vector, i. e. $\|v\|_2 = 1$. A rotation in $\mathbb{R}^3$ about the axis with direction vector v with the angle ϕ is given by following linear map:

$$x \longmapsto Ax = 2\begin{bmatrix} p_0^2 + p_1^2 - \frac{1}{2} & p_1\,p_2 - p_0\,p_3 & p_1\,p_3 + p_0\,p_2 \\ p_1\,p_2 + p_0\,p_3 & p_0^2 + p_2^2 - \frac{1}{2} & p_2\,p_3 - p_0\,p_1 \\ p_1\,p_3 - p_0\,p_2 & p_2\,p_3 + p_0\,p_1 & p_0^2 + p_3^2 - \frac{1}{2} \end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},$$

where

$$p_0 = \cos(\phi), \qquad \begin{bmatrix}p_1\\ p_2\\ p_3\end{bmatrix} = \sin(\phi)\,v.$$

Note: Here

$$\left\|\begin{bmatrix}p_0\\ p_1\\ p_2\\ p_3\end{bmatrix}\right\|_2 = 1.$$

I. e. the p-vector is a unit vector in a space of dimension 4.

The standard scalar product

We know the scalar product of two vectors $x = \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}, \; y = \begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$ is

$$x \cdot y = x^T y = x_1\,y_1 + \ldots + x_n\,y_n = \sum_{j=1}^{n} x_j\,y_j.$$

The Euclidean Length of a vector:

$$\|x\|_2 = |x| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{x^T x}. \qquad (1)$$

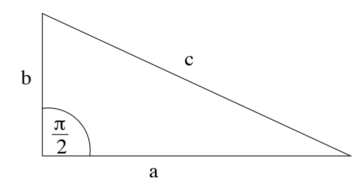
The angle between x and y:

$$\phi = \arccos\left(\frac{x^T y}{\|x\|_2\,\|y\|_2}\right). \qquad (2)$$

The equations (1) and (2) are definitions, motivated by elementary geometry.
(1) is Pythagoras's theorem.
(2) follows for vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ from the law of cosines (see also next slides).

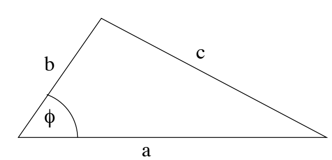
The law of cosines is a generalization of Pythagoras's theorem for arbitrary angles.

Pythagoras:



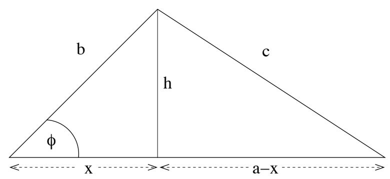
$$c^2 = a^2 + b^2$$

Cosinussatz:



$$c^2 = a^2 + b^2 - 2ab\cos\phi$$

Proof sketch of the law of cosines (here for acute angles only)



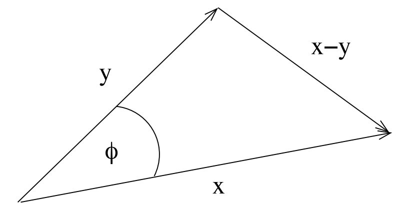
$$\begin{aligned} c^2 &= (a-x)^2 + h^2 \qquad\qquad b^2 = h^2 + x^2 \\ &= (a-x)^2 + b^2 - x^2 \\ &= a^2 - 2ax + x^2 + b^2 - x^2 \\ &= a^2 + b^2 - 2ax \end{aligned}$$

We have

$$x = b\cos\phi.$$

$$\Rightarrow \qquad c^2 = a^2 + b^2 - 2ab\cos\phi$$

Angle between two vectors and law of cosines.



With the law of cosines:

$$\|x-y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 - 2\,\|x\|_2\,\|y\|_2\cos\phi. \qquad (1)$$

We compute:

$$\begin{aligned} \|x-y\|_2^2 &= (x-y)^T(x-y) \\ &= (x^T - y^T)(x-y) \\ &= x^Tx - x^Ty - y^Tx + y^Ty \qquad\qquad (\text{with } x^Ty = y^Tx) \\ &= \|x\|_2^2 + \|y\|_2^2 - 2\,x^Ty \qquad (2) \end{aligned}$$

Comparing (1) and (2) gives

$$\|x\|_2\,\|y\|_2\cos\phi = x^Ty.$$

Then

$$\phi = \arccos\left(\frac{x^Ty}{\|x\|_2\,\|y\|_2}\right).$$

Perpendicular projection of a vector onto an axis.

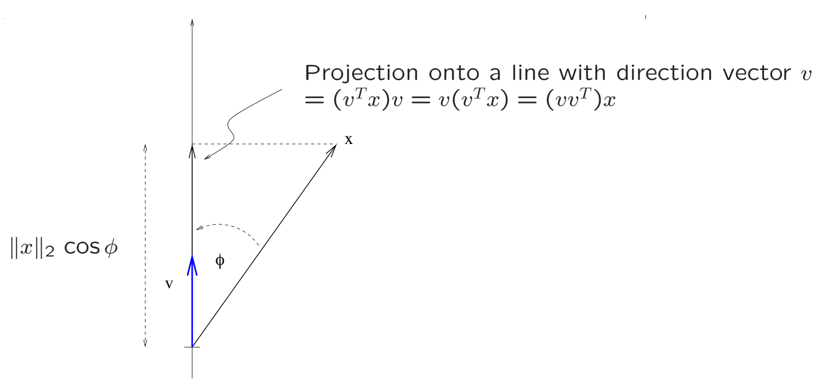
Recall: The angle ϕ between 2 vectors x, v is $\cos\phi = \frac{v^Tx}{\|v\|_2\,\|x\|_2}$.

If v has length 1, then we get $v^Tx = \|x\|_2\cos\phi$.

The perpendicular projection of x onto the line passing through the origin with direction vector v is then

$$x \longmapsto (v^Tx)\,v = v\,(v^Tx) = (v\,v^T)\,x.$$

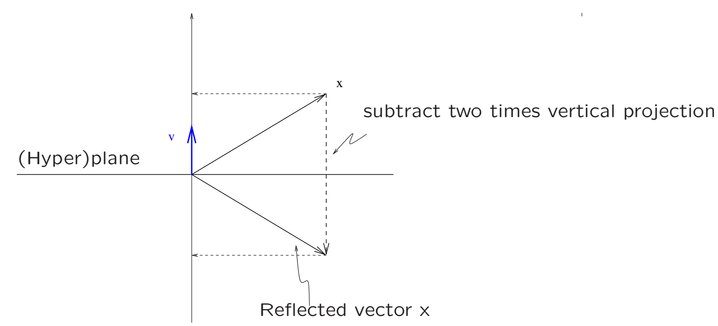
Conclusion: The projection is a (not invertible) linear map $x \mapsto Ax$ with the matrix $A = v\,v^T$.



Reflection at (hyper)plane

Let $v \in \mathbb{R}^n$ be a unit vector (i. e. $\|v\|_2 = 1$).

If we subtract from a vector x 2 times the perpendicular projection of x onto the line through v, then we reflected x at the (hyper)plane perpendicular to v.



This is again a linear map: $\qquad x \longmapsto x - 2\,vv^T\,x = \underbrace{(I - 2vv^T)}_{A}\,x.$

This reflection matrix A is usually called **Householder matrix**.

These matrices are very important for numerical analysis. More on this later.

Rotations and Reflections are length and angle preserving.

These linear transformations are called **orthogonal transformations**.

The corresponding matrices are called **orthogonal matrices**.

Orthogonal matrices

Theorem: Let $A \in \mathbb{R}^{n\times n}$. Then the following statements are equivalent:

1. $\|Ax\|_2 = \|x\|_2$ for all $x \in \mathbb{R}^n$.
2. $(Ax)^T(Ay) = x^Ty$ for all $x, y \in \mathbb{R}^n$.
3. $A^TA = I$ ($I=$ identity matrix)
4. $A^T = A^{-1}$.
5. The columns of A are perpendicular to each other and all have length 1. That means they form an **orthonormal basis** of $\mathbb{R}^n$. Formally:
$$A = [a_1 \ a_2 \ \dots \ a_n] \quad \Rightarrow \quad a_j^T a_k = \delta_{jk} = \begin{cases} 1 & \text{falls } k = j \\ 0 & \text{sonst} \end{cases}$$
6. The rows of A are perpendicular to each other and all have length 1.
7. $A\,A^T = I$.

If one of these conditions is satisfied, then all other conditions are also satisfied. In this case the matrix A is called **orthogonal**.

Orthogonal matrices satisfy $1 = \det(I) = \det(A^TA) = \det(A^T)\det(A) = \det(A)^2$, so
$$\det(A) = 1 \quad \text{oder} \quad \det(A) = -1.$$

Gramian matrix

Let $A = [a_1 \ a_2 \ \dots \ a_n] \in \mathbb{R}^{m\times n}$ an arbitrary matrix with columns a_k. Then
$$A^TA = \begin{bmatrix} a_1^Ta_1 & a_1^Ta_2 & \dots & a_1^Ta_n \\ a_2^Ta_1 & a_2^Ta_2 & \dots & a_2^Ta_n \\ \vdots & & & \vdots \\ a_n^Ta_1 & a_n^Ta_2 & \dots & a_n^Ta_n \end{bmatrix}.$$

The symmetric matrix A^TA has as entries all scalar products of the columns of A.

In particular

$A^TA = I$ ($I=$identity matrix) $\Leftrightarrow$ The columns of A are perpendicular and have length 1.

Definition: A^TA is called **Gramian matrix** of A.

Recall basic facts from Linear Algebra

Let $v_1, v_2, \dots, v_p$ be eigenvectors of $A \in \mathbb{C}^{n\times n}$, such that
$$Av_k = \lambda_k v_k, \qquad \lambda_k \in \mathbb{C}, \qquad k = 1, \dots, p. \qquad (*)$$
Then
$$AV = V\Lambda, \qquad (**)$$
where
$$V = [v_1 \ v_2 \ \dots v_p], \qquad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix}.$$

The converse is also true, i. e. $(*)$ follows from $(**)$ (see below).

If $p = n$ and the eigenvectors are linearly independent, i. e. they form a basis of $\mathbb{C}^n$, then the matrix V is invertible and then from $(**)$ it follows that
$$A = V\Lambda V^{-1}. \qquad (***)$$

This product representation is called **diagonalization** of A.

Not every matrix has a basis of eigenvectors
(keywords: generalized eigenvectors, Jordan normal form).

Derivation of $(**)$ from $(*)$: We compute
$$A\,[v_1 \ v_2 \ \dots v_p] = [Av_1 \ Av_2 \ \dots Av_p] = [\lambda_1 v_1 \ \lambda_2 v_2 \dots \lambda_p v_p] = [v_1 \ v_2 \ \dots v_p] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{bmatrix}.$$

For a matrix $A \in \mathbb{R}^{n\times n}$ we can distinguish the following cases :

- The matrix A does not have a basis of eigenvectors.
  Example: Let
$$\underbrace{\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}}_{J} \underbrace{\begin{bmatrix} \alpha \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}}_{v} = \lambda \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}, \qquad \alpha \in \mathbb{C}$$
  It can be shown that the Jordan matrix J does not have any eigenvectors beside the vectors v, i. e. a basis of eigenvectors does not exist.

- Eigenvalues and eigenvectors can be complex.
  Example rotation matrices:
$$\begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \begin{bmatrix} 1 \\ \pm i \end{bmatrix} = e^{\mp i\phi} \begin{bmatrix} 1 \\ \pm i \end{bmatrix}.$$

- The angle between two eigenvectors can be arbitrary and in particular very acute.
  Example:
$$\begin{bmatrix} 1 & x^2-1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \pm 1 \\ \pm 1 \end{bmatrix} = (\pm x) \begin{bmatrix} x \pm 1 \\ \pm 1 \end{bmatrix}$$
  The cosine of the angle ϕ_x between the two eigenvectors (scalar product divided by the product of the length of the two vectors) is $\lim_{x\to\infty} \cos(\phi_x) = 1$, i. e. the angle between the two eigenvectors converges to 0.

Theorem of principal axis and transformation of a symmetric matrix (spectral theorem):

Let $A \in \mathbb{R}^{n\times n}$ be a **symmetric** matrix. Then a diagonal matrix $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ and an orthogonal matrix $V = [v_1, \dots, v_n] \in \mathbb{R}^{n\times n}$ exist, such that
$$A = V\Lambda V^T.$$

The columns of V form an orthonormal basis of eigenvectors of A to the eigenvalues λ_k. More precise:
$$Av_k = \lambda_k v_k, \qquad \textit{and} \qquad v_j^T v_k = \delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

Note:
In $(*)$ we can write V^{-1} instead of V^T, because V is orthogonal, so we have $V^T = V^{-1}$.

Remarks on the spectral theorem

On the next pages we discuss some facts about symmetric matrices.
We derive them from the spectral theorem but they can also be proved directly.
These facts are consequences of the following identity

Fact 1: Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then
$$v_1^T A v_2 = v_2^T A v_1 \qquad \text{for all } v_1, v_2 \in \mathbb{R}^n.$$

Proof:
$$v_1^T A v_2 = v_1^T A^T v_2 = (Av_1)^T v_2 = v_2^T (Av_1) = v_2^T A v_1.$$

The left equation is true because A is symmetric.
The 3rd equation is true because $x^T y = \sum_k x_k y_k = y^T x$ for all $x, y \in \mathbb{R}^n$.

Fact 2: Let $A \in \mathbb{R}^{n\times n}$ be symmetric and let $v = v_1 + iv_2 \in \mathbb{C}^n$ be a complex vector with real part $v_1 \in \mathbb{R}^n$ and imaginary part $v_2 \in \mathbb{R}^n$. Then
$$\overline{v}^T A v \in \mathbb{R}.$$

Here $\overline{v} = v_1 - iv_2$ is the conjugate complex vector to v.

Proof:
$$\overline{v}^T A v = (v_1 - iv_2)^T A (v_1 + iv_2) = (v_1^T A v_1 - v_2^T A v_2) + i\,\underbrace{(v_1^T A v_2 - v_2^T A v_1)}_{=0 \text{(Fact 1)}} \in \mathbb{R}.$$

Fact 3: Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then every eigenvalue λ of A is real and has a corresponding real eigenvector.

Proof:
Let $Av = \lambda\, v$ with $\lambda \in \mathbb{C}$ and $v = v_1 + iv_2 \in \mathbb{C}^n \setminus \{0\}$, where v_1, v_2 real.
If we multiply the eigenvalue equation with $\overline{v}$ we get:

$$Av = \lambda\, v \qquad \Rightarrow \qquad \underbrace{\overline{v}^T Av}_{\in\mathbb{R}} = \lambda\, \underbrace{\overline{v}^T v}_{\in\mathbb{R}\setminus\{0\}} \qquad \Rightarrow \qquad \lambda \in \mathbb{R}.$$

Now we know that λ is real, so that splitting the equation $Av = \lambda\, v$ into real and imaginary part gives:

$$A(v_1 + iv_2) = \lambda\,(v_1 + iv_2) \qquad \Rightarrow \qquad Av_1 = \lambda\, v_1, \quad Av_2 = \lambda\, v_2.$$

Since $v \neq 0$, at least one of the real vectors v_1, v_2 is not 0, i. e. it is an eigenvector.

Fact 4: Let $A \in \mathbb{R}^{n\times n}$ symmetric and let $v_1, v_2 \in \mathbb{R}^n$ eigenvectors of A to distinct eigenvalues. Then $v_1^T v_2 = 0$, i. e. the eigenvectors are orthogonal.

Proof:

$$\begin{aligned} 0 &= v_2^T A v_1 - v_1^T A v_2 && \text{(Fact 1)} \\ &= v_2^T(\lambda_1 v_1) - v_1^T(\lambda_2 v_2) \\ &= \lambda_1\, v_1^T v_2 - \lambda_2\, v_2^T v_1 && \text{(We have } v_1^T v_2 = v_2^T v_1\text{)} \\ &= \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0}\, v_1^T v_2. \end{aligned}$$

$$\Rightarrow \qquad v_1^T v_2 = 0.$$

Eigenvalue criteria for positive definiteness

Theorem: Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then:
A is positive definite if and only if all eigenvalues of A are positive.

Proof of the eigenvalue criteria: With the spectral theorem an orthonormal-basis $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ of eigenvectors exist:

$$Av_j = \lambda_j\, v_j, \qquad v_j^T v_k = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

Then every vector $x \in \mathbb{R}^n$ can be written as linear combination of the eigenvectors:

$$x = x_1\, v_1 + x_2\, v_2 + \dots + x_n\, v_n, \quad x_j \in \mathbb{R}.$$

We compute:

$$\begin{aligned} x^T Ax &= (x_1\, v_1 + x_2\, v_2 + \dots + x_n\, v_n)^T A (x_1\, v_1 + x_2\, v_2 + \dots + x_n\, v_n) \\ &= (x_1\, v_1 + x_2\, v_2 + \dots + x_n\, v_n)^T (x_1\, \lambda_1\, v_1 + x_2\, \lambda_2\, v_2 + \dots + x_n\, \lambda_n\, v_n) \\ &= \lambda_1\, x_1^2\, \underbrace{v_1^T v_1}_{=1} + \lambda_2\, x_1 x_2\, \underbrace{v_1^T v_2}_{=0} + \dots \lambda_2\, x_2^2\, \underbrace{v_2^T v_2}_{=1} + \dots \\ &= \lambda_1\, x_1^2 + \lambda_2\, x_2^2 + \dots + \lambda_n\, x_n^2 \end{aligned}$$

If all eigenvalues are positive, then $x^T Ax > 0$ for all $x \neq 0$.
If one eigenvalue λ_k is negative, then $v_k^T A v_k = \lambda_k < 0$.

Gramian matrices are positive semi-definite

Let $A \in \mathbb{R}^{m\times n}$ an arbitrary real matrix. Then we have for all $x \in \mathbb{R}^n$,

$$x^T A^T A\, x = (Ax)^T (Ax) = \|Ax\|_2^2 \geq 0.$$

$\Rightarrow \qquad A^T A$ is positive semi-definite and therefore does not have negative eigenvalues.

If A has linear independent columns, then $Ax = 0$ only for $x = 0$.
Then $x^T A^T A x = 0$ only for $x = 0$.
Then $A^T A$ is positive definite and therefore has only positive eigenvalues.

Singular value decomposition (here only for square matrices)

The **singular value decomposition theorem** (short: SVD) can be formulated in 2 different ways:

Let $A \in \mathbb{R}^{n\times n}$ an arbitrary matrix.

Formulation 1: Two orthonormal basis $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ and $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ and non negative numbers $\sigma_1, \sigma_2, \ldots, \sigma_n \in \mathbb{R}$ exist, such that

$$Av_k = \sigma_k u_k, \qquad k = 1, \ldots, n \qquad (*)$$

Notation: The vectors u_k, v_k are called singular vectors.
The numbers σ_k are called singular values.

Formulation 2: Two orthogonal matrices $U, V \in \mathbb{R}^{n\times n}$ and a diagonal matrix Σ with non negative diagonal values exist, such that

$$A = U\,\Sigma\,V^T. \qquad (**)$$

Relation between both formulations: $(*)\ \Rightarrow$

$$A\,\underbrace{[v_1\ v_2\ \ldots\ v_n]}_{V} = [Av_1\ Av_2\ \ldots\ Av_n] = [\sigma_1 u_1\ \sigma_2 u_2\ \ldots\ \sigma_n u_n] = \underbrace{[u_1\ \ u_2\ \ \ldots\ \ u_n]}_{U}\underbrace{\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_n\end{bmatrix}}_{\Sigma}.$$

I. e.: $AV = U\Sigma$. Multiplication with $V^T = V^{-1}$ gives: $A = U\Sigma V^T$.

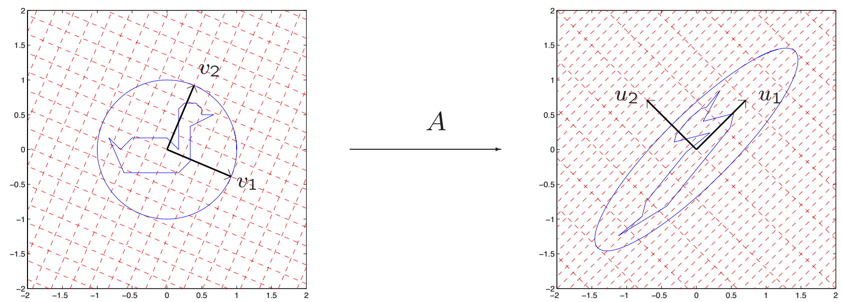
Illustration to formulation 1 of the singular value decomposition theorem.

The matrix $A \in \mathbb{R}^{2\times 2}$ in the example below has the singular values

$$\sigma_1 = 2, \qquad \sigma_2 = 1/2.$$

The singular vectors of A are v_1, v_2 and u_1, u_2. They form two orthonormal basis. We have

$$Av_1 = \sigma_1 u_1, \qquad Av_2 = \sigma_2 u_2.$$



Interpretation of the singular value decomposition $A = U\Sigma V^T$:

Every linear transformation $x \longmapsto Ax$ is a composition of

- an orthogonal (i. e. length and angle preserving) transformation V^T,
- an stretching/shortening Σ along the axis of a standard-coordinate system,
- an orthogonal (i. e. length and angle preserving) transformation U.

$$x \xrightarrow{V^T} V^T x \xrightarrow{\Sigma} \Sigma V^T x \xrightarrow{U} U\ \Sigma V^T x = Ax$$

See the examples on the next slide.

Conclusions from the singular value decomposition

Let $A = U\Sigma V^T$ be a singular value decomposition of A.

Then

$A^T = V\Sigma U^T$ is a singular value decomposition of A^T,
$A^{-1} = V\Sigma^{-1}U^T$ is a singular value decomposition of A^{-1} (if A^{-1} exist)

It also follows that

$$A^TA = (V\Sigma \underbrace{U^T)(U}_{I}\Sigma V^T) = V\Sigma^2 V^T, \qquad (*)$$

$$AA^T = (U\Sigma \underbrace{V^T)(V}_{I}\Sigma V^T) = U\Sigma^2 U^T. \qquad (**)$$

These are diagonalizations of symmetric and positive (semi)definite matrices A^TA, AA^T. And so the entries of

$$\Sigma^2 = diag(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)$$

are the eigenvalues of A^TA and AA^T. In other words:
The singular values of A are the square roots of the eigenvalues of A^TA and AA^T,

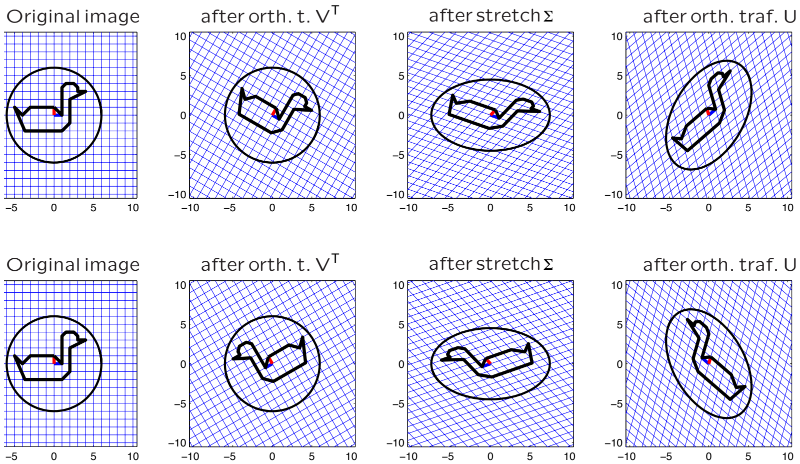
$$\sigma_k(A) = \sqrt{\lambda_k(A^TA)} = \sqrt{\lambda_k(AA^T)}.$$

From $(*)$ and $(**)$ we also get:

The singular vectors V are the eigenvectors of A^TA,
the singular vectors U are the eigenvectors of AA^T.

Example of a singular value decomposition, formulation 2

Singular value decomposition: $\mathbf{A} = \mathbf{U}\,\mathbf{\Sigma}\mathbf{V}^\mathbf{T}$ with $U^TU = I,\ V^TV = I,\ \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$



The Trafo V^T, Σ, U for $A = \begin{bmatrix} 0.87 & -0.25 \\ 0.75 & 0.87 \end{bmatrix}$ (above) and $A = \begin{bmatrix} -0.87 & 0.25 \\ 0.75 & 0.87 \end{bmatrix}$ (below).

In both cases $\sigma_1(A) \approx 1.25$, $\sigma_2(A) \approx 0.75$.

Computation of the singular value decomposition of an invertible matrix

The conclusions from the last pages lead to the following method to compute the singular value decomposition of an invertible matrix.

- Compute an orthonormal basis $V = [v_1\ \ v_2\ \ \ldots\ \ v_n]$ of eigenvectors of A^TA, such that $A^TA\,v_k = \lambda_k\,v_k$.
- The eigenvalues λ_k of the positive definite symmetric matrix A^TA are positive. Let $\sigma_k := \sqrt{\lambda_k}$. Then

$$(Av_j)^T(Av_k) = v_j^TA^TAv_k = v_j(\sigma_k^2\,v_k) = \sigma_k^2 v_j^T v_k = \sigma_k^2\,\delta_{jk}, \qquad (*)$$

  In particular $\|Av_k\|_2 = \sqrt{(Av_k)^T(Av_k)} = \sigma_k$.
- Let $u_k := Av_k/\sigma_k$ and form $U = [u_1\ \ u_2\ \ \ldots\ \ u_n]$. Then

$$Av_k = \sigma_k\,u_k \qquad (**)$$

  and from $(*)$ it follows, that the u_k form a orthonormal basis. With $(**)$ we have $AV = U\Sigma$ with $\Sigma := diag(\sigma_1, \ldots, \sigma_n)$, and so

$$A = U\Sigma V^T.$$

Singular value decomposition with Matlab

s = svd(A) gives back a vector with the singular values of A as entries.

[U,S,V] = svd(A) gives back the complete singular value decomposition: $A = USV^T$.

For symmetric matrices diagonalization and singular value decomposition are equal up to the sign

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and let

$$A = V\Lambda V^T, \qquad V^TV = I, \qquad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

an orthogonal diagonalization. Let

$$\Sigma = |\Lambda| = \mathrm{diag}(|\lambda_1|, \ldots, |\lambda_n|), \qquad D = \mathrm{diag}(\mathrm{sign}(\lambda_1), \ldots, \mathrm{sign}(\lambda_n)),$$

where sign$(\cdot)$ = is the sign. Then $\Lambda = D\,\Sigma$ and

$$A = \underbrace{V\,D}_{=:U}\Sigma V^T. \qquad (*)$$

U is orthogonal: $U^TU = (VD)^T(VD) = \underbrace{D^T}_{=D}\underbrace{V^TV}_{=I}D = D^2 = I.$

Therefore $(*)$ is a singular value decomposition.

If A is positive definite, then $D = I$ and $\Lambda = \Sigma$.

$\Rightarrow$ For positive definite symmetric matrices the eigenvalues are the singular values.

Singular value decomposition and condition numbers

From the singular value decomposition we can derive that

$$cond_2(A) = \|A\| \, \|A^{-1}\| = \frac{\sigma_{\mathsf{max}}(A)}{\sigma_{\mathsf{min}}(A)}.$$

More in the next lecture.