Lecture plan — Linear Algebra (401-0131-00L, HS25), ETH Zürich
Numbering of Sections, Definitions, Figures, etc. as in the Lecture Notes

Week 2

More on the span:
Corollary 1.27: Let $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_n \in \mathbb{R}^m$ such that $\mathbf{v}_k$ is a linear combination of the other vectors. Then

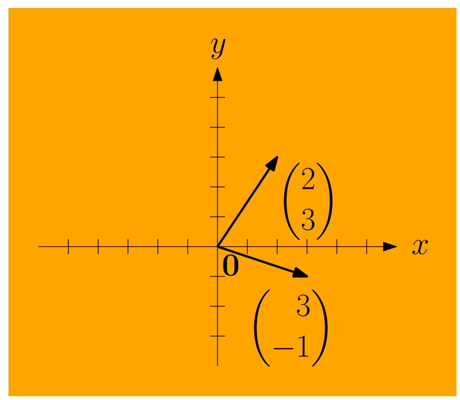
$$\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \underbrace{\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n)}_{S}.$$

Proof: By Lemma 1.26, S equals

$\underbrace{\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \mathbf{v}_{k+2}, \ldots, \mathbf{v}_n, \mathbf{v}_k)}_{T} = \mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ (order doesn't matter).

Challenge 1.6 in $\mathbb{R}^m$:
Lemma 1.28: Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^m$ be linearly independent. Then $\mathbf{Span}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m) = \mathbb{R}^m$.

This already captures the spirit of the Steinitz exchange Lemma 4.23. Proof in lecture notes and when we do Steinitz.



Matrices and linear combinations (Section 2.1)

Matrix: Notation for sequence of vectors:

$$\left(\begin{pmatrix}1\\3\\5\end{pmatrix}, \begin{pmatrix}2\\4\\6\end{pmatrix}\right) \quad \rightarrow \quad \begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} \quad \rightarrow \quad \left(\begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}3 & 4\end{pmatrix}, \begin{pmatrix}5 & 6\end{pmatrix}\right)$$

Also gives a sequence of covectors.
Definition 2.1: An $m \times n$ matrix is a table of real numbers with m rows and n columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

a_{ij}: entry in row i, column j
Dot-free notation: $A = [a_{ij}]_{i=1,j=1}^{m \ \ n}$
$\mathbb{R}^{m \times n}$: set of $m \times n$ matrices
Column notation (interprets columns as vectors):

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}, \quad \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^m$$

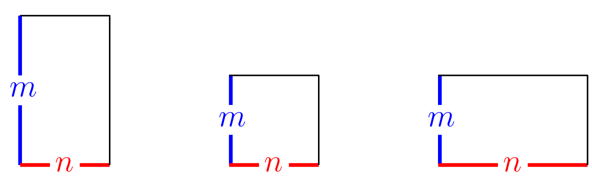
Row notation (interprets rows as covectors):

$$A = \begin{bmatrix} — & \mathbf{u}_1^\top & — \\ — & \mathbf{u}_2^\top & — \\ & \vdots & \\ — & \mathbf{u}_m^\top & — \end{bmatrix}, \quad \mathbf{u}_1^\top, \mathbf{u}_2^\top, \ldots, \mathbf{u}_m^\top \in (\mathbb{R}^n)^*$$

Matrix addition, matrix scalar multiplication: Definition 2.2

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}, \quad 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}.$$

0: zero matrix (all entries are 0)
Matrix shapes:



tall	square	wide
$m > n$	$m = n$	$m < n$

Square matrices: Definition 2.3

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 4 & 7 \\ 0 & 0 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 7 & 5 \end{bmatrix}$	$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 7 \\ 0 & 7 & 5 \end{bmatrix}$
identity I	diagonal	upper triangular	lower triangular	symmetric
$a_{ij} = \delta_{ij}$	$i \neq j : a_{ij} = 0$	$i > j : a_{ij} = 0$	$i < j : a_{ij} = 0$	$a_{ij} = a_{ji}$

Kronecker delta: $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.
$m \times m$ identity matrix $I = [\delta_{ij}]_{i=1,j=1}^{m \ \ m}$

Matrix-vector multiplication: Notation for linear combination of the columns

$$\underbrace{7\begin{pmatrix}1\\3\\5\end{pmatrix}+8\begin{pmatrix}2\\4\\6\end{pmatrix}}_{\text{linear combination}} \quad = \quad \underbrace{\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{pmatrix}7\\8\end{pmatrix}}_{\text{matrix-vector product}} \quad = \quad \underbrace{\begin{pmatrix}23\\53\\83\end{pmatrix}}_{\text{result}}$$

Definition 2.4: Let

$$A=\begin{bmatrix}| & | & & |\\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\\ | & | & & |\end{bmatrix}\in\mathbb{R}^{m\times n}, \quad \mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\mathbb{R}^n.$$

The vector

$$A\mathbf{x}:=\sum_{j=1}^{n}x_j\mathbf{v}_j\in\mathbb{R}^m$$

is the product of A and $\mathbf{x}$. "Transformation" of $\mathbf{x}\in\mathbb{R}^n$ to $A\mathbf{x}\in\mathbb{R}^m$

Observation 2.5:

(i) $\mathbf{b}\in\mathbb{R}^m$ is a linear comb. of the columns of A $\Leftrightarrow$ there is a vector $\mathbf{x}\in\mathbb{R}^n$ with $A\mathbf{x}=\mathbf{b}$

(ii) The columns of A are linearly independent $\Leftrightarrow$ $\mathbf{x}=\mathbf{0}$ is the only vector with $A\mathbf{x}=\mathbf{0}$.

Direct definition of $A\mathbf{x}$ (without columns): Observation 2.6

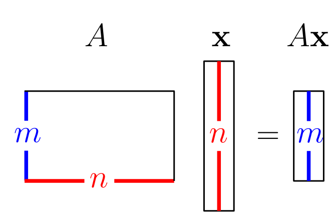
$$\begin{array}{c}\mathbf{u}_1^\top\\ \mathbf{u}_2^\top\\ \vdots\\ \mathbf{u}_m^\top\end{array}\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{bmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\vdots\\a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n\end{bmatrix}\begin{array}{c}\mathbf{u}_1^\top\mathbf{x}\\ \mathbf{u}_2^\top\mathbf{x}\\ \vdots\\ \mathbf{u}_m^\top\mathbf{x}\end{array}$$

$$\begin{array}{cccc}\mathbf{v}_1&\mathbf{v}_2&\cdots&\mathbf{v}_n\end{array}\qquad\qquad x_1\mathbf{v}_1\ +x_2\mathbf{v}_2\ +\cdots+x_n\mathbf{v}_n$$

Corollary 2.7: $I\mathbf{x}=\mathbf{x}$ for all $\mathbf{x}\in\mathbb{R}^m$.

Definition of $A\mathbf{x}$ in row notation: Observation 2.8

$$A=\begin{bmatrix}— & \mathbf{u}_1^\top & —\\ — & \mathbf{u}_2^\top & —\\ & \vdots & \\ — & \mathbf{u}_m^\top & —\end{bmatrix}, \quad A\mathbf{x}=\underbrace{\begin{pmatrix}\mathbf{u}_1^\top\mathbf{x}\\ \mathbf{u}_2^\top\mathbf{x}\\ \vdots\\ \mathbf{u}_m^\top\mathbf{x}\end{pmatrix}}_{\text{scalar products}}.$$

Pictorial view: 

Column space and rank:

Definition 2.9: Let A be an $m \times n$ matrix. The *column space* $\mathbf{C}(A)$ of A is the span of the columns,

$$\mathbf{C}(A) := \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m.$$

$\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{0} \in \mathbf{C}(A)$.

Fact 1.5:

$$\mathbf{C}\left(\begin{bmatrix} 2 & 3 \\ 3 & -1 \end{bmatrix}\right) = \mathbf{Span}\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}\right) = \mathbb{R}^2.$$

Independent columns:

Definition 2.10: Let

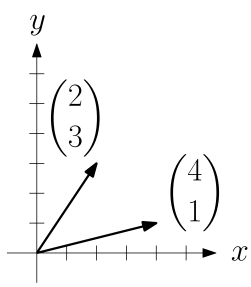
$$A = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}.$$

Column $\mathbf{v}_j$ is *independent* if $\mathbf{v}_j$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}$. Otherwise, $\mathbf{v}_j$ is *dependent*. The *rank* of A, $\mathbf{rank}(A)$, is the number of independent columns.

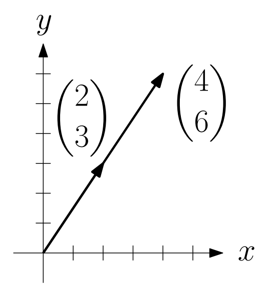
$\mathbf{rank}(A) = n$: linearly independent columns

$\mathbf{rank}(A) = 0$: zero matrix

$$\mathbf{rank}\left(\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}\right) = 2$$



$$\mathbf{rank}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right) = 1$$



both columns are independent

only first column is independent

Later: Reordering columns does not change the rank.

The independent columns span the column space:

Lemma 2.11: Let A be an $m \times n$ matrix with r independent columns, and let C be the $m \times r$ submatrix containing the independent columns. Then $\mathbf{C}(A) = \mathbf{C}(C)$.

Proof.

$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$: the independent columns.

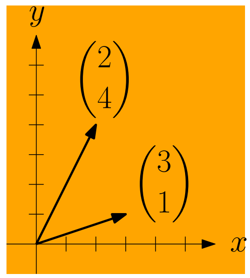
$\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-r}$: the dependent columns (order as in A)

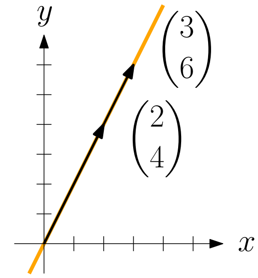
For all j, $\mathbf{w}_j$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \dots \mathbf{w}_{j-1}$: sequence contains all previous columns.

Take $\underbrace{\mathbf{Span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r)}_{\mathbf{C}(C)}$, add $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-r} \to \underbrace{\mathbf{Span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-r})}_{\mathbf{C}(A)}$.

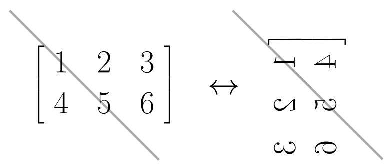
Adding linear combinations of previous vectors never changes the span (Lemma 1.26)! $\square$

Row space and transpose:
Row space: span of the rows (treated as columns of the *transposed* matrix)

$\mathbf{R}\left(\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix}\right)$ 

$\mathbf{R}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right)$ 

Transposition: Mirroring a matrix along the diagonal:



$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \leftrightarrow \quad A^\top = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Definition 2.12: Let $A = [a_{ij}]_{i=1,j=1}^{m\ \ n}$ be an $m \times n$ matrix. The *transpose* of A is the $n \times m$ matrix

$$A^\top := B = [b_{ij}]_{i=1,j=1}^{n\ \ m},$$

where $b_{ij} = a_{ji}$ for all i, j.

vector $\leftrightarrow$ covector $\qquad$ $m \times 1$ matrix $\leftrightarrow$ $1 \times m$ matrix

$$\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 & 5 \end{pmatrix} \qquad \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}^\top = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}$$

$(A^\top)^\top = A.$ $\qquad$ Observation 2.13

A symmetric $\Leftrightarrow A = A^\top$.

Definition 2.14: Let A be an $m \times n$ matrix. The *row space* $\mathbf{R}(A)$ of A is the column space of the transpose,

$$\mathbf{R}(A) := \mathbf{C}(A^\top) \subseteq \mathbb{R}^n.$$

Nullspace: everything that gets "transformed" to $\mathbf{0}$.

Definition 2.17: Let A be an $m \times n$ matrix. The *nullspace* of A is the set

$$\mathbf{N}(A) := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n.$$

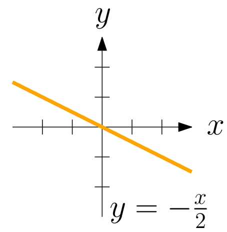
$\mathbf{0} \in \mathbf{N}(A)$ ($A\mathbf{0} = \mathbf{0}$).

$\mathbf{N}(A) = \{\mathbf{0}\} \Leftrightarrow$ columns of A are linearly independent (Observation 2.5 (ii)).

$\mathbf{N}(A) = \mathbb{R}^n \Leftrightarrow A = 0$.

$$\mathbf{N}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right) = ?\ \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad 2x + 4y = 0 \text{ and } 3x + 6y = 0 \quad \Leftrightarrow \quad y = -x/2.$$

$$\mathbf{N}\left(\begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}\right) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : y = -\frac{x}{2} \right\}$$



Matrices and linear transformation (Section 2.2)

Definition 2.18: Let A be an $m \times n$ matrix. The function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A : \underbrace{\mathbf{x}}_{\text{"input"} \in \mathbb{R}^n} \mapsto \underbrace{A\mathbf{x}}_{\text{"output"} \in \mathbb{R}^m}$$

is the *matrix transformation* with matrix A.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} : \qquad T_A : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

$$A = \begin{bmatrix} 1 & 1 \end{bmatrix} : \qquad T_A : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \end{pmatrix} \in \mathbb{R}^1$$

Lemma 2.19: "Linearity"
Let A be an $m \times n$ matrix, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ and $\lambda_1, \lambda_2 \in \mathbb{R}$. Then

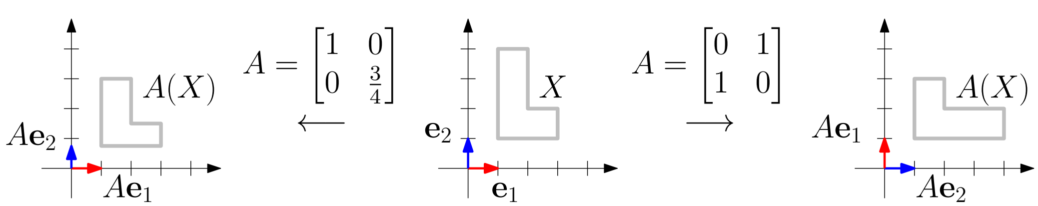
$$A(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 A\mathbf{x}_1 + \lambda_2 A\mathbf{x}_2.$$

Follows from the rules of vector addition (Definition 1.2), scalar multiplication (Definition 1.3) and matrix-vector multiplication (Definition 2.4).

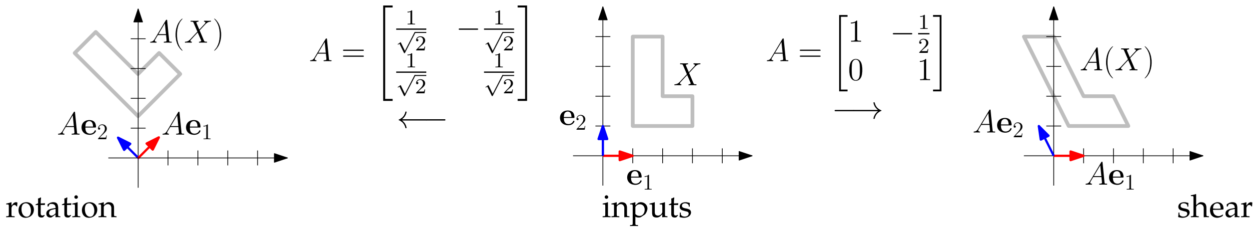
$$\begin{tikzcd}
\mathbf{x}_1, \mathbf{x}_2 \arrow[r, "\text{linear combination}"] \arrow[d, "T_A"'] & \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \arrow[d, "T_A"] \\
A\mathbf{x}_1, A\mathbf{x}_2 \arrow[r, "\text{linear combination}"'] & \lambda_1 A\mathbf{x}_1 + \lambda_2 A\mathbf{x}_2 = A(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2)
\end{tikzcd}$$

(commutative diagram)

Geometric interpretation: How is a set X of inputs transformed? $A(X) := \{A\mathbf{x} : \mathbf{x} \in X\}$
Examples $\mathbb{R}^2 \to \mathbb{R}^2$: $A\mathbf{e}_1, A\mathbf{e}_2$ are the two columns of A.



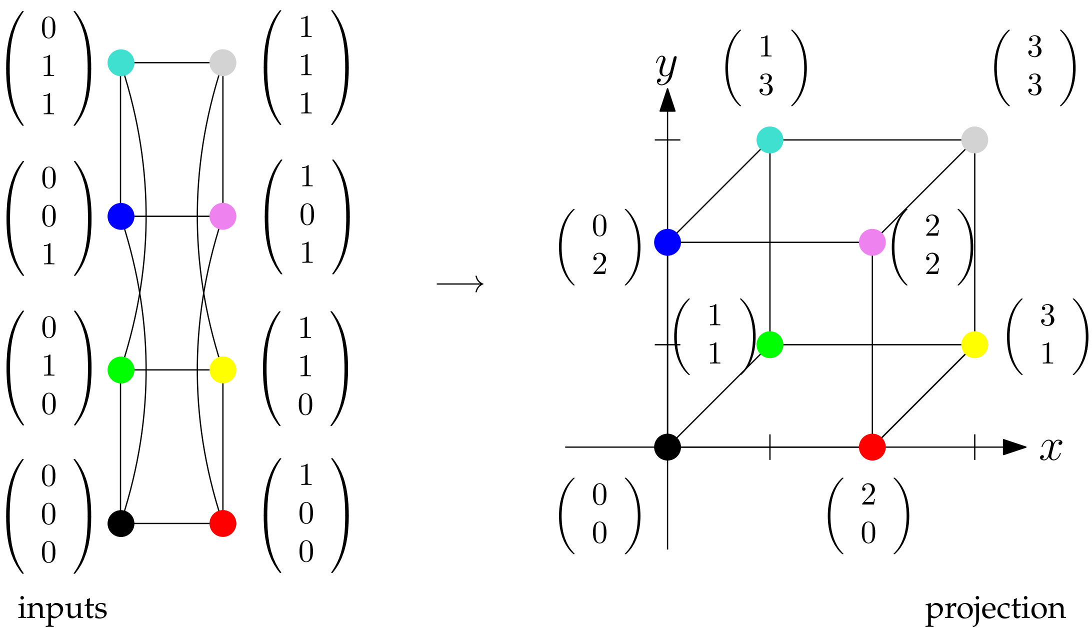
stretching inputs mirroring



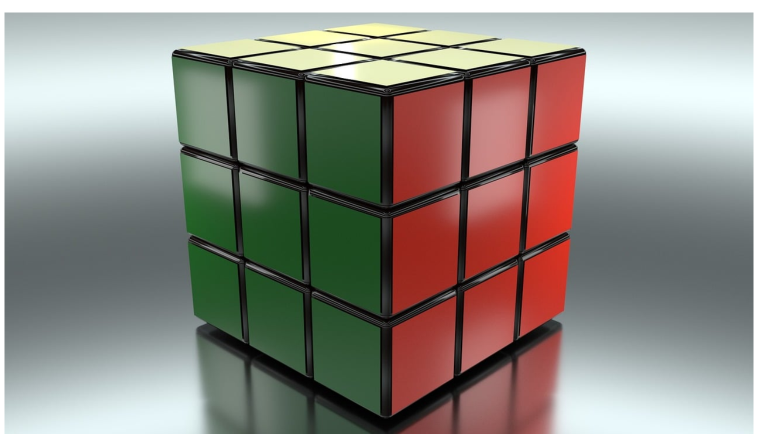
Animation: `https://scratch.mit.edu/projects/1076177705/`

Example $\mathbb{R}^3 \to \mathbb{R}^2$: *parallel projection*

$$\text{Applying } T_A : \mathbb{R}^3 \to \mathbb{R}^2 \text{ with } A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}:$$



Perspective projection:



Not a matrix transformation: lines that are parallel in the 3d-cube are not parallel in the projection (Exercise 2.20).