- Previous concept synopsis:
  - Vectors are groupings of scalars
  - Vector addition and multiplication of a vector by a scalar have been defined
  - Linear combination defined
  - Matrices are groupings of vectors
  - Matrix-vector product expresses a linear combination in a concise notation
- New concepts:
  - General definition of matrix
  - General definition of matrix-vector multiplication
  - Basic problems of linear algebra:
    - $\longrightarrow$ Solving linear systems
    - $\longrightarrow$ Least squares problems
    - $\longrightarrow$ Eigenvalue problems

**Definition.** *An $m$ by $n$ matrix is a grouping of $n$ vectors,*

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{pmatrix} \in S^{m \times n},$$

*where each vector has $m$ scalar components $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n \in S^m$.*

- Notation conventions:
  - scalars: normal face, Latin or Greek letters, $a, b, \alpha, \beta, u_1, a_{11}$, $A_{11}, I_{12}$
  - vectors: bold face, lower case Latin letters, $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{a}_1$
  - matrices: bold face, upper case Latin letters, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{L}_1$
- Matrix components

$$\boldsymbol{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \boldsymbol{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \boldsymbol{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \Rightarrow \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- A real-valued matrix with $m$ lines and $n$: $\boldsymbol{A} \in \mathbb{R}^{m \times n}$

- Instead of explicitly writing out components, it is often convenient to specify a matrix by a rule to construct each component

$$\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{A} = (a_{ij})$$

with indices taking values $i \in \{1, ..., m\}, j \in \{1, ..., n\}$.

Example: A *Hilbert matrix* $\boldsymbol{H}_m$ is defined as

$$\boldsymbol{H}_m \in \mathbb{Q}^{m \times m}, \boldsymbol{H}_m = \left( \frac{1}{i + j - 1} \right)$$

- Note that a vector is a matrix with a single column. The notation $\boldsymbol{v} \in \mathbb{R}^m$, is a customary shorter form of $\boldsymbol{v} \in \mathbb{R}^{m \times 1}$. The components of $\boldsymbol{v}$ can be specified as $\boldsymbol{v} = (v_i)$.

Example: A ones vector has all components equal to one

$$\boldsymbol{v} \in \mathbb{R}^m, \boldsymbol{v} = (1)$$

**Definition.** *Consider* $n$ *vectors, each with* $m$ *real components,* $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$*, grouped into the matrix* $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{pmatrix} \in \mathbb{R}^{m \times n}$*. Also consider* $n$ *real-valued scalars* $x_1$*,* $x_2, \ldots, x_n$*, grouped into a vector* $\boldsymbol{x} \in \mathbb{R}^n$*. The product of the* $m \times n$ *matrix* $\boldsymbol{A}$ *with the* $n \times 1$ *vector* $\boldsymbol{x}$ *is defined as the linear combination of the columns of* $\boldsymbol{A}$ *with coefficients given by components of* $\boldsymbol{x}$

$$\boldsymbol{A}\boldsymbol{x} = x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \ldots + x_n \boldsymbol{a}_n,$$

*and is an* $m$*-component vector,* $\boldsymbol{A}\boldsymbol{x} \in \mathbb{R}^m$*.*

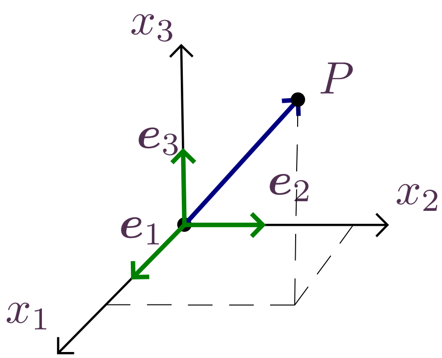
$$\boldsymbol{A}\boldsymbol{x} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \ldots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} x_1 a_{11} + x_2 a_{12} + \ldots + x_n a_{1n} \\ x_1 a_{21} + x_2 a_{22} + \ldots + x_n a_{2n} \\ \vdots \\ x_1 a_{m1} + x_2 a_{m2} + \ldots + x_n a_{mn} \end{pmatrix}$$

- In standard three-dimensional space, we often refer to the "x,y,z" directions, and state that to arrive at point with coordinates $(x,y,z)$ you travel distance $x$ along the "x" direction, $y$ along the "y" direction, $z$ along the "z" direction. Let us make this notion mathematically precise.

- Represent three-dimensional space as a set of vectors denoted as $\mathcal{E}_3$ ("Euclidean 3-space")

$$\mathcal{E}_3 = \left\{ \left( \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right), x_1, x_2, x_3 \in \mathbb{R} \right\}$$

Think of a vector $\boldsymbol{x}$ within $\mathcal{E}_3$ as starting at the origin and its tip at the point $P$ with coordinates $x_1, x_2, x_3$. This is the position vector of point $P$.




$$\boldsymbol{x} = x_1 \left( \begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right) + x_2 \left( \begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right) + x_3 \left( \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right) = \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right) \left( \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right) = \boldsymbol{I}\boldsymbol{x}$$

**Definition.** *A matrix is said to be square if it has the same number of lines and columns.*

**Definition.** *The diagonal of a square matrix* $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{m \times m}$ *is the vector* $\boldsymbol{d} = (d_i) \in \mathbb{R}^m$ *with components* $d_i = a_{ii}$*, with* $i = 1, 2, \ldots, m$*.*

**Definition.** *An identity matrix is a square matrix,* $\boldsymbol{I}_m = (I_{ij}) \in \mathbb{R}^{m \times m}$ *with components*

$$I_{ii} = 1 \text{ for } i = 1, 2, \ldots, m, I_{ij} = 0 \text{ for } i \neq j.$$

**Definition.** *The columns of the identity matrix* $\boldsymbol{I}_m \in \mathbb{R}^{m \times m}$ *are the canonical (or standard) basis vectors of* $\mathcal{V} = \mathbb{R}^m$*. The* $i^{\text{th}}$ *column vector is denoted by* $\boldsymbol{e}_i$

$$\boldsymbol{I}_m = (\ \boldsymbol{e}_1 \ \ \boldsymbol{e}_2 \ \ \ldots \ \ \boldsymbol{e}_m\ ) = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix}, \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \boldsymbol{e}_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

- Apply the definition of matrix-vector multiplication to product $\boldsymbol{I}_m\boldsymbol{b}$, with $\boldsymbol{I}_m \in \mathbb{R}^{m\times m}$ the identity matrix, and $\boldsymbol{b} = (b_i) \in \mathbb{R}^m$

$$\boldsymbol{I}_m\boldsymbol{b} = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + \ldots + b_m\boldsymbol{e}_m = b_1\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + b_m\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \boldsymbol{b},$$

  to observe that multiplication with the identity matrix leaves $\boldsymbol{b}$ unchanged

- The components $b_1, b_2, \ldots, b_m$ can be interpreted as the amount of displacements along $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m$ required to arrive at point $P = (b_1, b_2, \ldots, b_m) \in \mathbb{R}^m$, and are also referred to as the coordinates of point $P$.

- When the number of components is clear, the subscript on the identity matrix is suppressed, as in $\boldsymbol{b} \in \mathbb{R}^m$

$$\boldsymbol{I}\boldsymbol{b} = \boldsymbol{b},$$

  where from context it results that $\boldsymbol{I} \in \mathbb{R}^{m\times m}$.

- Let $\boldsymbol{A} = (\ \boldsymbol{a}_1 \ \ \boldsymbol{a}_2 \ \ \dots \ \ \boldsymbol{a}_m\ ) \in \mathbb{R}^{m \times n}, \boldsymbol{x} \in \mathbb{R}^n$. Interpret the result

$$\boldsymbol{b} = \boldsymbol{b}\boldsymbol{I} = \boldsymbol{A}\boldsymbol{x}$$

of the matrix vector multiplication as stating that the vector $\boldsymbol{b} \in \mathbb{R}^m$, can be obtained by two different linear combinations:

1. Linear combination of standard basis vectors

$$\boldsymbol{b} = b_1\boldsymbol{e}_1 + \ldots + b_m\boldsymbol{e}_m$$

2. Linear combination of $n$ column vectors of $\boldsymbol{A} = (\ \boldsymbol{a}_1 \ \ \dots \ \ \boldsymbol{a}_n\ )$,

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n.$$

- We can interpret a matrix-vector multiplication as providing the coordinates $(b_1, \ldots, b_m)$ of the point $P$ with position vector $\boldsymbol{b}$ in the standard basis vectors from knowledge of the coordinates of point $P$ expressed as a linear combination of the column vectors of $\boldsymbol{A}$.

1. *Solving a linear system*. Given the components of vector $\boldsymbol{b} \in \mathbb{R}^m$ in the standard basis vectors, what are the coordinates with respect to another set of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$?

$$\text{Given } \boldsymbol{b}, \text{ find } \boldsymbol{x} \text{ such that } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

   with $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ a matrix with column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$. Very often the matrix $\boldsymbol{A}$ is square $m = n$.

2. *Least squares problem*. Given the components of vector $\boldsymbol{b} \in \mathbb{R}^m$ in the standard basis vectors, find the linear combination of the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$ that is "as close as possible" to $\boldsymbol{b}$

$$\text{Given } \boldsymbol{b}, \text{ find } \boldsymbol{x} \text{ such that } \boldsymbol{A}\boldsymbol{x} \text{ is as close as possible to } \boldsymbol{b}.$$

3. *Eigenproblem*. Given a square matrix $\boldsymbol{A} \in \mathbb{C}^{m \times m}$ are there linear combinations that leave the direction of a matrix-vector product unchanged?

$$\text{Given } \boldsymbol{A} \in \mathbb{C}^{m \times m}, \text{ find } \boldsymbol{x} \in \mathbb{C}^m, \lambda \in \mathbb{C} \text{ such that } \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}.$$