- New concepts:
  - Matrix-matrix product
  - Matrix transpose
  - Transpose of matrix sums, products
  - Scalar product of two vectors

Definition. *Consider matrices* $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{pmatrix} \in \mathbb{R}^{m\times n}$, *and* $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_p \end{pmatrix} \in \mathbb{R}^{n\times p}$. *The matrix product* $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X}$ *is a matrix* $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_p \end{pmatrix} \in \mathbb{R}^{m\times p}$ *with column vectors given by the matrix vector products*

$$\boldsymbol{b}_k = \boldsymbol{A}\boldsymbol{x}_k, \text{for } k = 1, 2\dots, p.$$

- A matrix-matrix product is simply a set of matrix-vector products, and hence expresses multiple linear combinations in a concise way.
- The dimensions of the matrices must be compatible, the number of rows of $\boldsymbol{X}$ must equal the number of columns of $\boldsymbol{A}$.
- A matrix-vector product is a special case of a matrix-matrix product when $p = 1$.
- We often write $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{X}$ in terms of columns as

$$\begin{pmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_p \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_p \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}\boldsymbol{x}_1 & \dots & \boldsymbol{A}\boldsymbol{x}_p \end{pmatrix}$$

Definition. *Given a matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$,

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

the transpose of $\boldsymbol{A}$*, denoted as* $\boldsymbol{A}^T \in \mathbb{R}^{n \times m}$ *is*

$$\boldsymbol{A}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}.$$

Intuitively: “rows become columns, columns become rows”

- Recall that a vector is the special case of a matrix with a single column, $\boldsymbol{v} \in \mathbb{R}^{m \times 1}$. The transpose of a vector is $\boldsymbol{v}^T \in \mathbb{R}^{1 \times m}$ a matrix with a single row, known as a *row vector*.
- Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ expressed through its column vectors

$$\boldsymbol{A} = \left(\begin{array}{cccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{array} \right),$$

its transpose can be expressed as

$$\boldsymbol{A}^T = \left(\begin{array}{c} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \\ \vdots \\ \boldsymbol{a}_n^T \end{array} \right) \in \mathbb{R}^{n \times m}.$$

- For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$, $(\boldsymbol{u}+\boldsymbol{v})^T = \boldsymbol{u}^T + \boldsymbol{v}^T$. Proof: by direct computation

$$(\boldsymbol{u}+\boldsymbol{v})^T = \left(\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \right)^T = \begin{pmatrix} u_1+v_1 \\ \vdots \\ u_m+v_m \end{pmatrix}^T = \begin{pmatrix} u_1+v_1 & \dots & u_m+v_m \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & \dots & u_m \end{pmatrix} + \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} = \boldsymbol{u}^T + \boldsymbol{v}^T$$

- For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m \times n}$, $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$. Proof: by direct computation using column vectors of $\boldsymbol{A}, \boldsymbol{B}$

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_n \end{pmatrix}$$

$$(\boldsymbol{A}+\boldsymbol{B})^T = \begin{pmatrix} \boldsymbol{a}_1+\boldsymbol{b}_1 & \dots & \boldsymbol{a}_n+\boldsymbol{b}_n \end{pmatrix}^T = \begin{pmatrix} (\boldsymbol{a}_1+\boldsymbol{b}_1)^T \\ \vdots \\ (\boldsymbol{a}_n+\boldsymbol{b}_n)^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1^T+\boldsymbol{b}_1^T \\ \vdots \\ \boldsymbol{a}_n^T+\boldsymbol{b}_n^T \end{pmatrix} = \boldsymbol{A}^T + \boldsymbol{B}^T$$

- Consider $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{x} \in \mathbb{R}^n$. What is $(\boldsymbol{A}\boldsymbol{x})^T$? Recall that $\boldsymbol{A}\boldsymbol{x}$ is a linear combination of columns of $\boldsymbol{A}$

$$\boldsymbol{A}\boldsymbol{x} = \begin{pmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{pmatrix} \boldsymbol{x} = x_1 \boldsymbol{a}_1 + \dots + x_n \boldsymbol{a}_n$$

Take transpose of vector sum to obtain

$$(\boldsymbol{A}\boldsymbol{x})^T = x_1 \boldsymbol{a}_1^T + \dots + x_n \boldsymbol{a}_n^T = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_n^T \end{pmatrix} = \boldsymbol{x}^T \boldsymbol{A}^T \tag{1}$$

- Now consider $\boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{B} \in \mathbb{R}^{n \times p}$. Obtain $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T \boldsymbol{A}^T$ by applying (1)

$$(\boldsymbol{A}\boldsymbol{B})^T = \begin{pmatrix} \boldsymbol{A}\boldsymbol{b}_1 & \dots & \boldsymbol{A}\boldsymbol{b}_n \end{pmatrix}^T = \begin{pmatrix} (\boldsymbol{A}\boldsymbol{b}_1)^T \\ \vdots \\ (\boldsymbol{A}\boldsymbol{b}_n)^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1^T \boldsymbol{A}^T \\ \vdots \\ \boldsymbol{b}_n^T \boldsymbol{A}^T \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1^T \\ \vdots \\ \boldsymbol{b}_n^T \end{pmatrix} \boldsymbol{A}^T = \boldsymbol{B}^T \boldsymbol{A}^T$$

- A special case of a matrix-matrix product occurs when the two factors correspond to a row multiplying a column vector. The result is in this case a single scalar.

Definition. *Given two real-valued (column) vectors* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m$ *the scalar product is the matrix-matrix multiplication*

$$\boldsymbol{u}^T\boldsymbol{v} = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix} = \sum_{i=1}^{m} u_i v_i$$

Note the compatibility of number of components, $\boldsymbol{u} \in \mathbb{R}^{m\times 1} \Rightarrow \boldsymbol{u}^T \in \mathbb{R}^{1\times m}, \boldsymbol{v} \in \mathbb{R}^{m\times 1}$, $\boldsymbol{u}^T\boldsymbol{v} \in \mathbb{R}^{1\times 1} = \mathbb{R}$.

- The scalar product can be viewed as function taking two vectors as arguments and producing a single scalar as a result. The usual notation in this case is

$$\langle\, ,\rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \langle \boldsymbol{u}\, , \boldsymbol{v}\, \rangle = \boldsymbol{u}^T\boldsymbol{v} = \sum_{i=1}^{m} u_i v_i$$

with $\mathcal{V} = \mathbb{R}^m$.