Math 221 Lec 11

2.4: Elementary matrices, 2.5: Transpose

Asa Royal (ajr74)

October 3, 2023

1 Transpose

Remark. The transpose of a matrix A is the matrix $A^\intercal$ where $(A^\intercal)_{ij} = A_{ji}$. Entries are reflected over the diagonal

Remark. If $x, y \in \mathbb{R}^n$, we can think of their dot dot product $\mathbf{x} \cdot \mathbf{y}$ as

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\intercal \mathbf{y}$$

Remark. Note then that $\mathbf{a}^\intercal \mathbf{a} = \mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$

Remark. $\mathbf{x} \cdot \mathbf{y}$ produces a scalar, so $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^\intercal \mathbf{y}$ or $\mathbf{y}^\intercal \mathbf{x}$ (both of which are 1×1 matrices), but $\mathbf{x} \cdot \mathbf{y} \neq \mathbf{y}\mathbf{x}^\intercal$ or $\mathbf{x}\mathbf{y}^\intercal$, both of which would be $n \times n$ matrices.

Remark. The linear transformation represented by the $n \times n$ matrix $\mathbf{a}\mathbf{a}^\intercal$ is $\text{proj}_{\mathbf{a}}\mathbf{x}$. By expressing the projection formula in terms of $\mathbf{a}$ and $\mathbf{a}^\intercal$, we can clearly show that it is a function in terms of $\mathbf{a}$.

Proof.

$$\begin{aligned}
\text{proj}_{\mathbf{a}}\mathbf{x} &= \frac{\mathbf{x} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \\
&= \frac{\mathbf{a}^\intercal \mathbf{x}}{\|\mathbf{a}\|^2}\mathbf{a} && \text{(express dot product as matrix mult, per second remark above)} \\
&= \mathbf{a}\frac{\mathbf{a}^\intercal \mathbf{x}}{\|\mathbf{a}\|^2} && \text{(can move since one of the terms above is a scalar)} \\
&= \frac{\mathbf{a}\mathbf{a}^\intercal \mathbf{x}}{\mathbf{a} \cdot \mathbf{a}} \\
&= \left(\frac{\mathbf{a}\mathbf{a}^\intercal}{\mathbf{a}^\intercal \mathbf{a}}\right)\mathbf{x} && \text{(again, express dot product as matrix mult)}
\end{aligned}$$

□

Example (projection expressed with transposes).

$$\begin{aligned}
\text{proj}_{\begin{bmatrix}1\\1\\1\end{bmatrix}} &= \frac{\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}}{\begin{bmatrix}1\\1\\1\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix}} && \text{denom is } 1 \times n \text{ times } n \times 1 \text{, aka dot prod} \\
&= \frac{1}{3}\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix}
\end{aligned}$$

Definition 1 (orthogonal matrix). An $n \times n$ matrix A is orthogonal if $A^\intercal A = I_n$, which is true iff $\mathbf{a}_i \cdot \mathbf{a}_j = 0$ for $i \neq j$ and 1 for $i = j$.

2 Left multiplication (row vector * matrix)

Remark. Let A be an $n \times m$ matrix. Let $\mathbf{x}$ be a vector of length n. The operation in which a row vector is multiplied by a matrix can be expressed as $\mathbf{x}^\intercal \mathbf{A}$.

Proposition 2. $A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot \mathbf{A}^\intercal \cdot \mathbf{y}$ (To move the matrix across a dot product operator, we must transpose it)

Proof.

$$\begin{aligned}
(A\mathbf{x}) \cdot \mathbf{y} &= (A\mathbf{x})^\intercal \mathbf{y} && \text{dot product as matrix mult} \\
&= \mathbf{x}^\intercal A^\intercal \mathbf{y} \\
&= \mathbf{x}^\intercal (A^\intercal \mathbf{y}) \\
&= \mathbf{x} \cdot A^\intercal \mathbf{y}
\end{aligned}$$

□

3 Elementary matrices

Proposition 3. Every invertible matrix can be expressed as a product of elementary matrices.

Proof. Since A is invertible, we can row reduce it to the identity matrix. We do this by multiplying A on the left by a series of elementary matrices $E = (E_k)(\ldots)(E_2)(E_1)$ such that $EA = I$. Elementary matrices are invertible, so we can multiply both sides of that equation by E^{-1}. Then $A = E^{-1} = ({E_1}^{-1})({E_2}^{-1})(\ldots)({E_k}^{-1})$. □

Remark. When we apply elementary operations to the rows of a matrix A, we multiply A on the left by an elementary matrix E, such that we get a transformed version of the rows of A.

$$EA = \begin{bmatrix} — & E_1 A & — \\ — & E_2 A & — \\ & \vdots & \\ — & E_m A & — \end{bmatrix}$$

Remark. We construct a row swap elementary matrix E by taking I_m and interchanging rows i and j to swap A_i and A_j. A row of E, E_k, looks like

$$\begin{cases} {e_k}^\intercal, & \text{if } k \neq i \text{ or } j \\ {e_j}^\intercal, & \text{if } k = i \\ {e_i}^\intercal, & \text{if } k = j \end{cases}$$

where e_i is the i-th basis vector and ${e_i}^\intercal$ is the i-th basis row vector.
Thus,

$$E_k A = \begin{cases} A_k, & \text{if } k \neq i \text{ or } j \\ A_j, & \text{if } k = i \\ A_i, & \text{if } k = j \end{cases}$$