Matrix calculations

Walter Mudzimbabwe

Matrix operations

Let $A \in \mathbb{R}^{m\times n}$ then we write

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

<u>Addition</u>: $\mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{m\times n}$: $\quad C = A + B$ where $c_{ij} = a_{ij} + b_{ij}$

<u>Scalar multiplication</u>: $\mathbb{R} \times \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{m\times n}$: $\quad C = \alpha A$ where $c_{ij} = \alpha a_{ij}$

<u>Multiplication</u>: $\mathbb{R}^{m\times n} \times \mathbb{R}^{n\times p} \longrightarrow \mathbb{R}^{m\times p}$: $\quad C = AB$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

<u>Transpose</u>: $\mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{n\times m}$: $\quad C = A^T$ where $c_{ij} = a_{ji}$

Special square matrices

Symmetric matrix: An $A \in \mathbb{R}^{n\times n}$ is symmetric if $A^T = A$.
Examples:

$$\begin{bmatrix} 1 & 7 & 3 \\ 7 & 4 & 5 \\ 3 & 5 & 0 \end{bmatrix}, \quad \begin{bmatrix} 4 & 0 & 1 & 10 \\ 0 & -3 & 6 & -2 \\ 1 & 6 & 1 & 10 \\ 10 & -2 & 10 & 10 \end{bmatrix}.$$

Orthogonal matrix: An $A \in \mathbb{R}^{n\times n}$ is orthogonal if $A^TA = I_n$ where $I_n \in \mathbb{R}^{n\times n}$ is an identity matrix.
The columns and rows are orthonormal ie if you take the dot product of any two columnns or rows, the product is 0.
Examples:

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Orthogonal matrices

From $A^T A = I_n$ we have $A^{-1} = A^T$ which makes finding the inverse much easier (than doing lots of row operations, imagine if $n = 1000$!)

Other Special non square matrices

Diagonal matrix: An $A \in \mathbb{R}^{m\times n}$ is a diagonal matrix if $a_{ij} = 0$ when $i \neq j$.
Notation: We write $A = \text{diag}(\alpha_1, \alpha_2, \cdots, \alpha_k)$ where $k = \min\{m, n\}$ then

$$A = [a_{ij}] \text{ is diagonal and } a_{ii} = \alpha_i \text{ for } i = 1, 2, \cdots, k$$

Examples:

$$\begin{bmatrix} 7 & 0 \\ 0 & -5 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \end{bmatrix}, \quad \begin{bmatrix} 4 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -13 \end{bmatrix}.$$

Not that this is not square.

Vector norms

A vector norm on $\mathbb{R}^n$ is a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that:

1. $f(x) \geq 0, \qquad \forall\, x \in \mathbb{R}^n$
2. $f(x+y) = f(x) + f(y), \qquad \forall\, x, y \in \mathbb{R}^n$
3. $f(\alpha x) = \alpha f(x), \qquad \forall\, \alpha \in \mathbb{R}, \forall\, x \in \mathbb{R}^n.$

We denote such an $f(x)$ by $||x||$.

Examples:

Holder/p-norms: $||x||_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$ For example:

1. $||x||_1 = |x_1| + |x_2| + \cdots + |x_n|$
2. $||x||_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{1/2} = (x^T x)^{1/2}$
3. $||x||_\infty = \max_i |x_i|$

Vector norms

Exercise: Prove that the 2-norm is invariant under orthogonal transformations

Solution: By a an orthogonal transformation of x, we mean Qx where Q is an orthogonal matrix. So the question is asking you to prove that $||Qx||_2 = ||x||_2$.

Now

$$
\begin{aligned}
||Qx||_2^2 &= (Qx)^T Qx \\
&= x^T Q^T Qx \\
&= x^T x, \text{ since } Q^T Q = I_n, \text{ because } Q \text{ is orthogonal} \\
&= ||x||_2^2
\end{aligned}
$$

which implies $||Qx||_2 = ||x||_2$ since the norm can not be negative.

Matrix norms

A matrix norm on $\mathbb{R}^{m\times n}$ is a function $f : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}$ such that:

1. $f(A) \geq 0, \qquad \forall\, A \in \mathbb{R}^{m\times n}$ with $f(A) = 0$ iff $A = 0$
2. $f(A + B) = f(A) + f(B), \qquad \forall\, A, B \in \mathbb{R}^{m\times n}$
3. $f(\alpha A) = \alpha f(A), \qquad \forall\, \alpha \in \mathbb{R}, \forall\, A \in \mathbb{R}^{m\times n}$.

Again, we denote such an $f(x)$ by $||x||$.

Examples:

Frobenius norm: $$||A||_F = \left(\sum_i^m \sum_j^n |a_{ij}|^2\right)^{1/2}.$$

We can also define p-norms but they are not neccessary in this course.

Singular value decomposition (SVD)

Let $A \in \mathbb{R}^{m \times n}$ then there exist orthogonal matrices

$$U = [u_1, u_2, \cdots, \cdots, u_m] \in \mathbb{R}^{m \times m}$$
$$V = [v_1, v_2, \cdots, v_n] \in \mathbb{R}^{n \times n}$$

such that

$$U^T A V = \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_p) \tag{1}$$

where $p = \min\{m, n\}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$.
We can write (1) as

$$A = U \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_p) V^T$$

which is called the singular decomposition (SVD) of A.
The σ_i's are called singular values of A and vectors u_i and v_i are the i^{th} left and right singular vectors respectively.

Singular value decomposition (SVD)

We can also verify that

$$
\begin{aligned}
\mathrm{A}\mathrm{v}_i &= \sigma_i \mathrm{u}_i \\
\mathrm{A}^T\mathrm{u}_i &= \sigma_i \mathrm{v}_i
\end{aligned}
$$

To do this we need to verify that

$$\mathrm{A} = \sum_{j=1}^{r} \sigma_j \mathrm{u}_j \mathrm{v}_j^T$$

which implies

$$\mathrm{A}^T = \sum_{j=1}^{r} \sigma_j \mathrm{v}_j \mathrm{u}_j^T$$

Singular value decomposition (SVD)

Therefore

$$
\begin{aligned}
\mathrm{Av}_i &= \left(\sum_{j=1}^{r} \sigma_j \mathrm{u}_j \mathrm{v}_j^T\right) \mathrm{v}_i \\
&= \sum_{j=1}^{r} \sigma_j \mathrm{u}_j \mathrm{v}_j^T \mathrm{v}_i \\
&= \sigma_i \mathrm{u}_i \mathrm{I}_n \\
&= \sigma_i \mathrm{u}_i
\end{aligned}
$$