Ramp-up Mathematics — Analysis
Homework Sheet 2

Exercise 2.1
Show that the operator norm of $A \in \mathbb{R}^{m\times n}$ with respect to the ℓ^∞-norm in the domain and range is the row-sum norm from Example 3.4.2.

Proof. This follows from

$$\frac{\|Ax\|_\infty}{\|x\|_\infty} = \frac{\max_j |\sum_\ell A_{j\ell}x_\ell|}{\|x\|_\infty} \leq \max_j \sum_\ell |A_{j\ell}|,$$

with equality for $x = (x_j)_{j=1,\dots,n}$ with $x_\ell = \operatorname{sgn}(A_{j^*,\ell})$, where j^* is such that $\max_j \sum_\ell |A_{j\ell}| = \sum_\ell |A_{j^*\ell}|$ the maximum is attained. □

Exercise 2.2
Let $A \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$. Show the identities

$$x^T Ax = \operatorname{trace}(xx^T A),$$
$$\|x\|_2^2 = \operatorname{trace}(xx^T).$$

Hint: Use that the trace is cyclic, i.e., $\operatorname{trace}(ABC) = \operatorname{trace}(CAB)$ (if the dimensions fit).

Proof.
1. We have $(xx^T)_{j\ell} = x_j x_\ell$ and so $[(xx^T)A]_{jk} = \sum_\ell x_j x_\ell A_{\ell k}$ and $\operatorname{Tr}(xx^T A) = \sum_{j,\ell} x_j x_\ell A_{\ell,j} = \langle x, Ax\rangle = x^T Ax$, as desired.
2. We have $(xx^T)_{j\ell} = x_j x_\ell$ and so $\operatorname{Tr}(xx^T) = \sum_j x_j^2 = \|x\|_2^2$.

□

Let us also give the proof of the ciclicity of the trace. More precisely, for $A \in \mathbb{R}^{n\times n}$, define the trace on the $\mathbb{R}^{n\times n}$-matrices by

$$\operatorname{Tr}_n(A) := \sum_{j=1}^n A_{jj}.$$

Lemma 0.1. *Let $m, n \in \mathbb{N}$, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$. Then*

$$\operatorname{Tr}_m(AB) = \operatorname{Tr}_n(BA).$$

Proof. This follows from

$$\operatorname{Tr}_m(AB) = \sum_{j=1}^m (AB)_{jj} = \sum_{j=1}^m \sum_{\ell=1}^n A_{j\ell}B_{\ell j} = \sum_{\ell=1}^n \sum_{j=1}^m B_{\ell j}A_{j\ell} = \sum_{\ell=1}^n (BA)_{\ell\ell} = \operatorname{Tr}_n(BA),$$

as desired. □

Exercise 2.3
Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be given by $f(A) = A^3 - A + A^T$. What is $Df(A)[H]$ for some $H \in \mathbb{R}^{n\times n}$?

Proof. Expanding $(A+H)^3 - (A+H) + (A+H)^T$ shows that

$$Df(A)[H] = A^2H + AHA + HA^2 - H + H^T.$$

□

Exercise 2.4
Let $f(A, x) = Ax$ for $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$. What is the derivative of f with respect to x? What is the derivative with respect to A? (Let's denote the former by $D_x f(A, x)$ and the latter by $D_A f(A, x)$.)

Proof.

1. For $h \in \mathbb{R}^n$, we expand $f(A, x+h) = Ax + Ah$. Thus, $D_x f(A, x) = A$ as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$.

2. For $H \in \mathbb{R}^{m\times n}$, we expand $f(A+H, x) = Ax + Hx$. Thus, $D_x f(A, x)[H] = Hx$ as a linear map from $\mathbb{R}^{m\times n}$ to $\mathbb{R}^m$.

□

Exercise 2.5
Let $B \in \mathbb{R}^{n\times n}$ and $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = x^T Bx$. What is $Df(x)$ and what is $\nabla f(x)$?

Proof. We have $f(x+h) = x^T Bx + h^T Bh + h^T Bx + x^T Bh$. Observe $h^T Bx = \langle h, Bx\rangle = \langle Bx, h\rangle = (Bx)^T h$. Hence, $Df(x) = (Bx)^T + x^T B$ as a map from $\mathbb{R}^n$ to $\mathbb{R}$, i.e., a $1 \times n$-matrix. Analogously, $\nabla f(x) = Bx + B^T x$ as a map from $\mathbb{R}$ to $\mathbb{R}^n$, i.e., a $n \times 1$-matrix. □