Question

Let $Q : \mathbf{R}^n \to \mathbf{R}$ be a quadratic form given by $Q(x) = x^T Ax$ where A is an $n \times n$ symmetric matrix. Show from the definition of differentiation that its derivative at the point $x \in \mathbf{R}^n$ is the $1 \times n$ matrix $2x^T A$. What does this amount to when $n = 1$? Let M_n denote the space of $n \times n$ (real) matrices, which can be thought of as $\mathbf{R}^{n^2}$. Let $f : M_n \to M_n$ be the function given by

$$f(A) = (A^T A).$$

Show from the definition that $df(A)H = 2(A^T H)$ for every $H \in M_n$. What does this amount to when $n = 1$?

Answer

$$\begin{aligned}
Q(x) &= x^T Ax \\
Q(x+h) &= (x+h)^T A(x+h) \\
&= x^T Ax + h^T Ax + x^T Ah + h^T Ah \\
&= Q(x) + \underbrace{2x^T Ah}_{\textbf{linear in } h} + \underbrace{h^T Ah}_{\textbf{quadratic}} \\
\textbf{as } \underbrace{h^t Ax = x^T A^T h}_{\textbf{scalar}} &= x^T Ah
\end{aligned}$$

So we read off the linear term to see $dQ(x)h = 2x^T Ah$.
When $n = 1$ this simply says $\frac{d}{dx}(ax^2) = 2ax$.

$$\begin{aligned}
f(A) &= \textbf{trace}(A^T A) \\
f(A+H) &= \textbf{trace}(A+H)^T (A+H) \\
&= \textbf{trace}(A^T A) + \textbf{trace}(H^T A) + \textbf{trace}(A^T H) + \textbf{trace}(H^T H) \\
&= f(A) + \underbrace{2\textbf{trace}(A^T H)}_{\textbf{linear in } H} + \underbrace{\textbf{trace}(H^T H)}_{\textbf{quadratic}}
\end{aligned}$$

So we read of the linear to see $df(A)H = 2\textbf{trace}(A^T H)$.
When $n = 1$ this says $\frac{d}{da}(a^2) = 2a$.