Trace Duality

Claim: If $f \in C^1(I)$, and X, H are self adjoint matrices, then there is a unique quantity $g(X)$ such that

$$\mathrm{tr} Df(X)[H] = \mathrm{tr} Hg(X).$$

We start with a construction from Bhatia's Matrix Analysis: Let $f \in C^1(I)$ and define $f^{[1]}$ on $I \times I$ by

$$f(\lambda, \mu) = \begin{cases} \frac{f(\lambda)-f(\mu)}{\lambda-\mu} & \lambda \neq \mu \\ f'(\lambda) & \lambda = \mu. \end{cases}$$

We call $f^{[1]}(\lambda, \mu)$ the *first divided difference* of f at (λ, μ). If Λ is a diagonal matrix with entries $\{\lambda_i\}$, We may extend f to accept Λ by defining the (i,j)-entry of $f^{[1]}(\Lambda)$ to be $f^{[1]}(\lambda_i, \lambda_j)$. If A is a self adjoint matrix with $A = U\Lambda U^*$, then we define $f^{[1]}(A) = Uf^{[1]}(\Lambda)U^*$. Now we borrow a theorem from Bhatia:

Theorem 1 (Bhatia V.3.3)**.** *Let $f \in C^1(I)$ and let A be a self adjoint matrix with all eigenvalues in I. Then*

$$Df(A)[H] = f^{[1]}(A) \circ H,$$

where $\circ$ denotes the Schur-product in a basis where A is diagonal.

That is, if $A = U\Lambda U^*$, then

$$Df(A)[H] = U \left(f^{[1]}(\Lambda) \circ (U^*HU) \right) U^*.$$

To prove our claim, we need to take the trace of both sides. Since trace is invariant under a change of basis, it is clear that

$$\mathrm{tr} Df(A)[H] = \mathrm{tr} \left(f^{[1]}(\Lambda) \circ (U^*HU) \right).$$

If $U = u_{ij}$, $U^* = \overline{u}_{ij}$ and $H = h_{ij}$, then the (i,j)-entry of U^* is

$$(U^*HU)_{ij} = \overline{u}_{ik} h_{k\ell} u_{\ell j}$$

Where we sum over the duplicate indices k and ℓ. While the structure of $f^{[1]}(\Lambda)$ is a bit unruly, our diagonal entries are $f'(\lambda)$. This means that when we take the trace of the Schur product, we have

$$\sum_i f(\lambda_i) \overline{u}_{ik} h_{k\ell} u_{\ell i}$$

Now consider the matrix product $U \,\mathrm{diag}\{f'(\lambda_1), \ldots, f'(\lambda_n)\}\, U^*H$. Since one of our terms is diagonal, the trace of this multiplication is simple:

$$\mathrm{tr}\, U \,\mathrm{diag}\{f'(\lambda_1), \ldots, f'(\lambda_n)\}\, U^*H = \sum_k u_{ik} f(\lambda_k) \overline{u}_{k\ell} h_{\ell i}$$

Since our entries commute, we can relable our indices $i \mapsto \ell$ $\ell \mapsto k$ $k \mapsto i$ to get

$$\mathrm{tr}\, U \,\mathrm{diag}\{f'(\lambda_1), \ldots, f'(\lambda_n)\}\, U^*H = \sum_i f(\lambda_i) \overline{u}_{ik} h_{k\ell} u_{\ell i},$$

See that the quantity a $U\,\mathrm{diag}\{f'(\lambda_1), \ldots, f'(\lambda_n)\}\,U^*H$ is independent of our choice of H, so is it the needed quantity $g(X)$. Further, since $X = U\Lambda U^*$, $g(X) = f'(X)$. This recovers theorem 3.3 of *LEARN TO DO CITATIONS* as we have

$$\mathrm{tr}\ Df(X)[H] = \mathrm{tr} Hg(X)$$