MATH 3210 HOMEWORK #2

ALLEN KNUTSON

Q. Compute the derivative of the determinant map $M_n(\mathbb{R}) \to \mathbb{R}$.

A. If M is invertible,

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\det(M+\varepsilon A) - \det(M)) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\det(M(1+\varepsilon M^{-1}A)) - \det(M)) \\
&= \lim_{\varepsilon\to 0} \det(M)\frac{1}{\varepsilon}(\det(1+\varepsilon M^{-1}A) - 1) \\
&= \lim_{\varepsilon\to 0} \det(M)\frac{1}{\varepsilon}(1+\varepsilon\,\mathrm{trace}(M^{-1}A) + O(\varepsilon^2)) \\
&= \det(M)\mathrm{trace}(M^{-1}A) \\
&= \mathrm{trace}\left(\det(M)M^{-1}A\right) \\
&= \mathrm{trace}\left(\mathrm{adj}(M)A\right)
\end{aligned}$$

where $\mathrm{adj}(M)$ is the numerator of Cramer's rule for the inverse of a matrix.

Note that the formula for $\mathrm{adj}(M)$ makes sense (and is continuous) even when M is not invertible. Since the invertible matrices are dense in all matrices, this continuous formula that holds for invertible matrices must also hold for noninvertible.

https://en.wikipedia.org/wiki/Cramer%27s_rule

https://en.wikipedia.org/wiki/Adjugate_matrix

Q. Compute the critical points (in the source) and critical values (in the target) of the map $f : M \mapsto MM^T$, from $M_n(R) \to$ symmetric matrices.

HINT: it's easier to compute the perpendicular space to the image than the image itself. Two matrices A,B are perpendicular if $\mathrm{trace}(AB^T) = 0$.

A. The derivative is $Df_M(A) = AM^T + MA^T$, as we computed in class. If its image is not full rank, it's because the image is perpendicular to S, i.e. $\mathrm{trace}((AM^T + MA^T)S) = 0$ for all matrices A. Rewriting,

$$\begin{aligned}
0 = \mathrm{trace}((AM^T + MA^T)S) &= \mathrm{trace}(AM^TS) + \mathrm{trace}(MA^TS) \\
&= \mathrm{trace}(AM^TS) + \mathrm{trace}(S^TAM^T) \qquad \text{since } \mathrm{trace}(X) = \mathrm{trace}(X^T) \\
&= \mathrm{trace}(AM^TS) + \mathrm{trace}(AM^TS^T) \qquad \text{since } \mathrm{trace}(XY) = \mathrm{trace}(YX) \\
&= \mathrm{trace}(A(M^TS + M^TS^T)) \\
&= \mathrm{trace}(A(M^TS + M^TS)) \qquad \text{since } S = S^T \\
&= 2\mathrm{trace}(AM^TS)
\end{aligned}$$

The only vector (in the vector space of matrices!) M^TS perpendicular to all A is the zero vector, i.e. $M^TS = 0$.

If M is invertible, then the only such S is zero, i.e. M is a regular point. If M is not invertible, then we can pick $\vec{v} \neq \vec{0}$ such that $M^T\vec{v} = \vec{0}$, and let $S = \vec{v}(\vec{v})^T$, so M is not a regular point.

If M is invertible, then so is M^T and MM^T. If M is not invertible, then neither is MM^T. So the critical values are the noninvertible symmetric matrices.

Alternate answer: we're mapping from an n^2-dimensional vector space to an $\binom{n+1}{2}$-dimensional one. So by the nullity plus rank theorem, the derivative is onto iff its kernel is only $\binom{n}{2}$-dimensional, not larger. What is that kernel?

$$\begin{aligned} A \in \ker(Df_M) &\iff AM^T + MA^T = 0 \\ &\iff (AM^T) + (AM^T)^T = 0 \\ &\iff AM^T \text{ is antisymmetric} \end{aligned}$$

If M (hence M^T) is invertible, then such A correspond $1:1$ to antisymmetric matrices, so the nullity is indeed only $\binom{n}{2}$-dimensional. It is more annoying to show that if M is not invertible, then the space of A such that AM^T is antisymmetric is larger than $\binom{n}{2}$-dimensional.

Q. Let $O(n)$ denote the set of orthogonal matrices, i.e. $f^{-1}(\mathbf{1})$. Compute the tangent space to each point of $O(n)$.

A. Let $M \in O(n)$. Then

$$T_M O(n) = \ker(Df_M) \quad = \quad \{A \,:\, Df_M(A) = 0\} = \{A \,:\, AM^T + MA^T = 0\}$$

In particular, at $M = \mathbf{1}$, A must be antisymmetric. More generally, $(AM^T)^T = MA^T = -AM^T$, i.e. AM^T (also known as AM^{-1}) must be antisymmetric.

Q. Let G be an $n \times n$ diagonal matrix with k consecutive 1s and $n-k$ consecutive 0s on the diagonal. Let $h : O(n) \to$ symmetric matrices take $M \mapsto MGM^T$.
Compute its derivative and the rank of the derivative.

A. $Dh_M(A) = AGM^T + MGA^T$ much as before. If $M = \mathbf{1}$, the kernel is

$$\left\{A \,:\, A + A^T = 0, AG + GA^T = 0\right\}.$$

If we write $G = \begin{pmatrix} \mathbf{1}_k & 0 \\ 0 & \mathbf{1}_{n-k} \end{pmatrix}$ and $A = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$, then these say $P = -P^T$, $S = -S^T$, $Q = R = 0$. So the kernel is $\binom{k}{2} + \binom{n-k}{2}$ dimensional, and the rank is $\binom{n}{2} - \binom{k}{2} - \binom{n-k}{2} = k(n-k)$.

For general $M \in O(n)$, we want

$$\left\{B \,:\, BM^T + MB^T = 0, BGM^T + MGB^T = 0\right\}$$

Multiply each equation by M^T, M on left and right, to obtain

$$\left\{B \,:\, M^TB + B^TM = 0, M^TBG + GB^TM = 0\right\}$$

Now under the correspondence $A = M^T B$, these are the same equations as before, i.e. the kernel has the same dimension as before. Hence the rank is again $k(n-k)$, independent of $M \in O(n)$.

E-mail address: `allenk@math.cornell.edu`