Math 25a Solution Set #9 (Part B)

Isidora Milin

December 2001

Problem 1

Let $O(n) \subset Mat(n,n)$ be the set of orthogonal matrices, i.e., matrices whose columns form an orthonormal basis of $\mathbb{R}^n$. Let $S(n,n)$ be the space of symmetric $n \times n$ matrices, and $A(n,n)$ be the space of antisymmetric $n \times n$ matrices.

Part (a). We want to show: $A \in O(n) \Leftrightarrow A^T A = I$. If we let A_k be the k-th column of A, and a_{ij} the i,j-th entry of A, we get:

$$\begin{aligned}
A^T A &= \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \\
&= \begin{pmatrix} A_1 \cdot A_1 & A_1 \cdot A_2 & \dots & A_1 \cdot A_n \\ A_2 \cdot A_1 & A_2 \cdot A_2 & \dots & A_2 \cdot A_n \\ \vdots & \vdots & \ddots & \dots \\ A_n \cdot A_1 & A_n \cdot A_2 & \dots & A_n \cdot A_n \end{pmatrix}
\end{aligned}$$

Using the multiplication above, we get that:

$A \in O(n)$

$\Leftrightarrow$ columns of A are an orthon. basis of $\mathbb{R}^n$

$\Leftrightarrow A_i \cdot A_j = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{for } i \neq j. \end{cases}$

$\Leftrightarrow A^T A = I$ ∎

Part (b). We first show that $A,B \in O(n) \Rightarrow AB \in O(n)$. Using the equivalence shown in part (a), we have that $AB \in O(n) \Leftrightarrow (AB)^T AB = I$ and the claim follows easily: $A \in O(n) \Rightarrow A^T A = I$, $B \in O(n) \Rightarrow B^T B = I$.

$$(AB)^T AB = (B^T A^T)AB = B^T (A^T A)B = B^T IB = B^T B = I$$

The implication $A \in O(n) \Rightarrow A^{-1} \in O(n)$ is also simple, given part (a). First we need to check that all $A \in O(n)$ are invertible, but this follows easily, either from the fact that A has n LI columns (which was shown to be sufficient for invertibility on a previous homework), or from the fact that, by (a), it has a left inverse - namely - A^T. We know that the existence of left inverse guarantees invertibility, and, (surprise!) the inverse is actually the left inverse, so that we conclude:

$$A \in O(n) \Rightarrow A^{-1} = A^T \Rightarrow I = AA^T = (A^T)^T A^T \Rightarrow A^{-1} = A^T \in O(n)$$

∎

Part (c). To show that $A^T A - I$ is symmetric, we do not have to consider it entry by entry; showing that it equals its own transpose is sufficient (and faster):

$$(A^T A - I)^T = (A^T A)^T - I^T = A^T (A^T)^T - I = A^T A - I$$

Things you need to know (for general purposes) about matrix transpose include: $(AB)^T = B^T A^T$, $(A + B)^T = A^T + B^T$, $(A^T)^T) = A$ and $(A^T)^{-1} = (A^{-1})^T$.

Part (d). Here a map $F : Mat(n,n) \to S(n,n)$ was defined as $F(A) = A^T A - I$. By part (a), we have that $F(A) = 0 \Leftrightarrow A^T A = I \Leftrightarrow A \in O(n)$, so that $O(n) = F^{-1}(0)$. By part (c), we know that $im(F) \subset S(n,n)$.We want to show the following:

$$\forall A \text{ invertible } [DF(A)] : Mat(n,n) \to S(n,n) \text{ is onto.}$$

We begin by determining the linear map $DF(A)$. We claim that $\forall H \in Mat(n,n)$ we have $DF(A)(H) = A^T H + H^T A$. It is obviously a linear map, but it may not be so obvious why it is the derivative. This becomes clear as we go about to prove that - one way to get what the derivative is is to look at $F(A + H) - F(A)$ and disregard terms not depending linearly on H, for they will converge to 0 faster:

$$\begin{aligned}
& lim_{|H| \to 0} \frac{|F(A+H) - F(A) - (A^T H + H^T A)|}{|H|} \\
= \;& lim_{|H| \to 0} \frac{|(A+H)^T (A+H) - I - A^T A - (A^T H) - (H^T A)|}{|H|} \\
= \;& lim_{|H| \to 0} \frac{|A^T A + H^T A + A^T H + H^T H - A^T A - A^T H - H^T A|}{|H|} \\
= \;& lim_{|H| \to 0} \frac{|H^T H|}{|H|} \\
= \;& 0
\end{aligned}$$

The last step follows from the Cauchy-Shwarz inequality:

$$0 \le \frac{|H^T H|}{|H|} \le \frac{|H^T||H|}{|H|} = |H^T|$$

Since $|H| \to 0 \Rightarrow |H^T| \to 0$ (if a norm of a matrix approaches 0, all of its entries approach 0 too, so the same holds for H^T, since it has the same entries as H). Thus $DF(A) : Mat(n,n) \to S(n,n)$ such that $H \mapsto A^T H + (A^T H)^T$.

We shall prove that $DF(A)$ is surjective by showing that $dim(im(DF(A))) = dim(S(n,n))$, which entails calculating both dimensions in terms of n.

First, notice that $S(n,n)$ is a subspace of the n^2-dimensional vector space $Mat(n,n)$. If E_{ij} is the standard basis for $Mat(n,n)$ where $i,j \in 1,\ldots,n$, then $B = \{ E_{ii} \mid i = 1,\ldots,n \} \cup \{ E_{ij} + E_{ji} \mid i \neq j \}$ is a basis for $S(n,n)$. It is clear that elements of B are LI - no two of them have a non-zero entry at the same position. They span $S(n,n)$ – any symmetric matrix is a linear combination with its entries as scalars. Since B has $n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}$ elements, we conclude that $dim(S(n,n)) = \frac{n(n+1)}{2}$.

To calculate $dim(im(DF(A)))$ we shall employ the rank-nullity theorem, and for that reason we first examine $dim(ker(DF(A)))$. Note that

$$Ker(DF(A)) = \{ H \mid A^T H = -H^T A = -(A^T H)^T \} = \{ H \mid A^T H \in A(n,n) \}$$

Now we can easily construct a bijection f: $Ker(DF(A)) \to A(n,n)$, $H \mapsto A^T H$. This is a linear map whose inverse is: f^{-1}: $A(n,n) \to Ker(DF(A))$, $H \mapsto (\mathrm{A}^T)^{-1} H$. We know that f^{-1} exists because A is invertible, so that A^T is as well. Now that we have constructed a bijective linear map f, we conclude that $dim(Ker(DF(A))) = dim(A(n,n))$.

The basis for $A(n,n)$ can be found similarly as for $S(n,n)$, in terms of the standard basis for $Mat(n,n)$. It is $\{ E_{ij} - E_{ji} \mid j < i \}$ and thus has cardinality $\frac{n(n-1)}{2}$, given that that is the number of possible pairs i,j such that $j < i$. We conclude that:

$$dim(Ker(DF(A))) = dim(A(n,n)) = \frac{n(n-1)}{2}$$

Now the Rank-nullity theorem applied to the linear map $DF(A)$ gives us:

$$dim(Ker(DF(A))) + dim(Im(DF(A))) = dim(Mat(n,n)$$

$$dim(Im(DF(A))) = n^2 - \frac{n(n-1)}{2} = \frac{n(n+1)}{2} = dim(S(n,n))$$

Since we know that $Im(DF(A))$ is a subspace of $S(n,n)$, we conclude $Im(DF(A)) = S(n,n)$ and $DF(A)$ is surjective.

Note that this is an example of a map which is surjective but not injective - $DF(A)$ has a nontrivial kernel, but since it is a map between two spaces of different dimension, this does not prevent it from being surjective. A mistake a lot of people made was to assume that to prove the map is surjective they need to show it has a trivial kernel, which is false in this case! ∎

Part (e). To show that $O(n)$ is a manifold embedded in $Mat(n,n)$, recall, from part (d), that $O(n) = F^{-1}(0)$. Now use Thm 3.2.3 from the textbook, which gives you a description of a manifold by equations, to conclude $O(n)$ is one. The continuous map $F : Mat(n, n) \to S(n, n)$ is such that $O(n) \subset Mat(n, n)$ is a solution to the equation $F(x) = 0$.

As for showing $T_I O(n) = A(n, n)$, just recall that the tangent space at a point of the manifold M is the kernel of the linear map which is the value of the derivative of the equation of M at that point:

$$\begin{aligned}
T_I O(n) &= Ker(DF(I)) \\
&= \{ H \in Mat(n, n) \mid I^T H = -(I^T H)^T \} \\
&= \{ H \mid H = -H^T \} \\
&= A(n, n)
\end{aligned}$$

∎