Math 395 IBL - The Spectral Theorem

William Garland, Michael Ivanitskiy, Le Hu

Friday October 13, 2017

Context: Finite Dimensional Spectral Theorem

Let H be some symmetric matrix. Then, $\exists U \in O(n)$ such that

$$H = UXU^{-1} \qquad U \in O(n)$$

where X is a diagonal matrix; i.e.

$$X = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Take an input H. We desire some $U \in O(n)$ for which $U^{-1}HU$ is diagonal. Define a function $f : M_n \to \mathbb{R}$ by

$$f(z) = \sum r_i z_{ii}$$

for $r_i \in \mathbb{R}$ with the property that $r_1 < r_2 < \cdots < r_n$ and let $g : M_n \to \mathbb{R}$ be defined by

$$g(U) = f(U^{-1}HU)$$

Example: Let

$$f(Z) = z_{11} + 2z_{22} \qquad H = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \qquad U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in O(n)$$

Then

$$\begin{aligned} U^{-1}HU &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta + 2\sin\theta & 2\cos\theta + 3\sin\theta \\ -\sin\theta + 2\cos\theta & -2\sin\theta + 3\cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\theta + 4\cos\theta\sin\theta + 3\sin^2\theta \\ \sin^2\theta - 4\cos\theta\sin\theta + 3\cos^2\theta \end{bmatrix} \\ f &= \left(\cos^2\theta + 4\cos\theta\sin\theta + 3\sin^2\theta\right) + 2\cdot\left(\sin^2\theta - 4\cos\theta\sin\theta + 3\cos^2\theta\right) \end{aligned}$$

This function is minimized for a value of θ satisfying

$$\begin{aligned}
0 &= \frac{d}{d\theta}\left(\left(\cos^2\theta + 4\cos\theta\sin\theta + 3\sin^2\theta\right) + 2\left(\sin^2\theta - 4\cos\theta\sin\theta + 3\cos^2\theta\right)\right) \\
&= -2\cos\theta\sin\theta + 4\cos(2\theta) + 6\sin\theta\cos\theta + 4\sin\theta\cos\theta - 8\cos(2\theta) - 12\cos\theta\sin\theta \\
&= -4\sin\theta\cos\theta - 4\cos(2\theta) \\
&= -2(\sin(2\theta) + 2\cos(2\theta)) \\
&\implies \tan(2\theta) = -2 \implies \theta \approx 1.01722
\end{aligned}$$

We wish to show that for any such symmetric matrix H, $\exists U \in O(n)$ such that $f(U^{-1}HU)$ is minimized, and furthermore that in this case $U^{-1}HU$ is diagonal.

Problem 19: Show that $\exists\ U_0 \in O(n) \ni g$ is minimized at some U_0 on $O(n)$.

Proof. Recall that $g : O(n) \to \mathbb{R}$ is a continuous real-valued map. Therefore, it is sufficient to show that $O(n)$ is compact. Since $O(n)$ is a subset of a real vector space $\mathbb{R}^{n^2}$, we may show that it is closed and bounded.

Closed: Define a function $h : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ (M_n denotes the space of $n \times n$ matrices) by $h(A) = A^tA$. Note matrix multiplication is a polynomial map with respect to the entries of A, so h is continuous. Furthermore

$$O(n) = \{A \in M_n : A^tA = \mathrm{Id}\} = \{A \in M_n : h(A) = \mathrm{Id}\} = h^{-1}(\mathrm{Id})$$

Since h is continuous and the singleton $\{\mathrm{Id}\}$ is closed, we conclude that $O(n)$ is closed.

Bounded: Let $|\cdot|_\infty$ denote the supremum norm on M_n. Let $U \in O(n)$. Note that by the standard inner product on $\mathbb{R}^n$, we know that U is orthogonal implies $(Ue_1, Ue_2, \ldots, Ue_n)$ forms an orthonormal basis of $\mathbb{R}^n$. Furthermore note that these are the columns of U, and that if each has $|Ue_i| = 1$ under the 2-norm, then no element of U has absolute value > 1. Therefore $|U|_\infty \leq |U_{ij}| \leq 1$ for some i, j, so U is bounded by 1 for all $U \in O(n)$. Therefore $O(n)$ is bounded.

Thus $O(n)$ is compact. The image of a compact set under a continuous map is compact, so we note that $g(O(n)) \subset \mathbb{R}$ achieves its minimum at some $U_0 \in O(n)$. $\square$

We showed on a previous homework that if g achieves its minimum at U_0, then $[Dg]_{U_0}[v] = 0$ for all $v \in T_{U_0}O(n)$. But for any $U \in O(n)$, how do we describe T_UO_n? Recall that since $U \in O(n)$ which is a smooth manifold, then the function $U\exp : P \to B$ (where P is an open subset of $\mathfrak{so}(n)$) is a coordinate patch from $\mathfrak{so}(n)$ to $O(n)$ near U. Therefore,

$$T_UO(n) = [D(U\exp)]_U[\mathfrak{so}(n)] = [D(U)]_{\mathrm{Id}}[D\exp]_0[\mathfrak{so}(n)]$$

Note U as a map from $A \to UA$ is linear, so $[D(U)]_{\mathrm{Id}} = U\mathrm{Id} = U$. As we have previously shown, $[D\exp]_0 = \mathrm{Id}$, so consequently we conclude

$$T_UO(n) = U\mathrm{Id}[\mathfrak{so}(n)] = U \cdot \mathfrak{so}(n)$$

Thus if $A \in T_U O(n)$, then $A = UJ$ for some $J \in \mathfrak{so}\,(n)$.

However, we can also define $T_U O(n)$ in a different way, as the set of derivatives of paths through $O(n)$ at U. Under this definition, if $UJ \in T_U$, let $\gamma(t) : (-\delta, \delta) \to O(n)$ by $\gamma(t) = U\exp(tJ)$ and note $\gamma(0) = U\exp(0) = U$ and $\gamma'(0) = UJ\exp(0) = UJ$. By the Chain Rule, it follows that

$$\begin{aligned}[Dg]_U[UJ] &= \frac{d}{dt}\bigg|_{t=0} g(\gamma(t)) \\ &= \frac{d}{dt}\bigg|_{t=0} f((U\exp(tJ))^{-1}H(U\exp(tJ))) \\ &= \frac{d}{dt}\bigg|_{t=0} f(\exp(tJ)^{-1}U_0^{-1}HU\exp(tJ)) \\ &= \frac{d}{dt}\bigg|_{t=0} f(\exp(-tJ)X\exp(tJ))\end{aligned}$$

where $X = U^{-1}HU$ does not depend on J or t. Since f is linear, $[Df] = f$, so we conclude

$$[Dg]_U[UJ] = f\left(\frac{d}{dt}\bigg|_{t=0} \exp(-tJ)X\exp(tJ)\right) = f(-JX + XJ)$$

for any $UJ \in T_U O(n)$.

Now that we have computed $[Dg]$, we are ready to tackle Problem 20.

Problem 20: Show $U^{-1}HU$ is diagonal if and only if $g(U)$ is minimized.

First, suppose $U \in O(n)$ such that $X := U^{-1}HU$ is diagonal. Therefore $\forall J \in \mathfrak{so}\,(n)\,(n)$, the diagonal entries of J must all be zero, so note

$$-JX + XJ = \begin{bmatrix} 0 & & * \\ & \ddots & \\ -* & & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} + \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} 0 & & -* \\ & \ddots & \\ * & & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & & & & \\ \lambda_1 * & 0 & & & \\ \vdots & \lambda_2 * & 0 & & \\ \vdots & \vdots & & \ddots & \\ \vdots & \vdots & & & 0 \end{bmatrix} + \begin{bmatrix} 0 & \lambda_1 * & \cdots & \cdots & \\ & 0 & \lambda_2 * & \cdots & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \star & \star \\ * & \ddots & \star \\ * & * & 0 \end{bmatrix}$$

So

$$[Dg]_U\,[UJ] = \sum r_k \left[-JX_0 + X_0 J\right]_k k = 0$$

for all $UJ \in T_U O(n)$. Thus $g(U)$ is minimized.

Now, let $U \in O(n)$ and define $X = U^{-1}HU$ such that

$$[Dg]_U[UJ] = \sum r_k(-JX + XJ)_{kk} = 0$$

We need to show that X is diagonal.

Notice

$$X^t = (U^{-1}HU)^t = U^t H^t (U^{-1})^t = U^{-1}HU = X$$

so X is symmetric. Suppose X is not diagonal. Then $\exists i, j$ with $i \neq j$ such that

$$X_{ij} = X_{ji} = a \neq 0$$

Define $J \in \mathfrak{so}(n)$ such that $J_{ij} = 1, \quad J_{ji} = -1$, and $J_{rs} = 0$ for all $(r, s) \neq (i, j)$. Then note

$$JX_{ii} = J_{ij}X_{ji} = X_{ji} = a \qquad JX_{jj} = J_{ji}X_{ij} = -X_{ij} = -a$$

$$XJ_{ii} = X_{ij}J_{ji} = -X_{ij} = -a \qquad XJ_{jj} = X_{ji}J_{ij} = X_{ji} = a$$

and for $r \neq i, j$,

$$JX_{rr} = XJ_{rr} = 0$$

Therefore,

$$\begin{aligned} 0 = f(-JX + XJ) &= \sum r_k(-JX + XJ)_{kk} \\ &= r_i(-JX_{ii} + XJ_{ii}) + r_j(-JX_{jj} + XJ_{jj}) \\ &= r_i(-a + -a) + r_j(a + a) = 2(r_j - r_i)a \end{aligned}$$

But $r_j \neq r_i$ by construction (since $j \neq i$), so it follows that $a = 0$, a contradiction. Therefore it must be the case that $X = U^{-1}HU$ is diagonal.

Problem 21: Prove the Spectral Theorem.

The result is immediate: if H is a symmetric matrix, then by Problem 19 $\exists U_0 \in O(n)$ such that $g(U_0)$ is minimized. Define $X_0 = U_0^{-1}HU_0$. Since g attains a local minimum on $O(n)$ at U_0, it follows that $\forall U_0 J \in T_{U_0}O(n)$,

$$[Dg]_{U_0}[U_0 J] = f(-JX_0 + X_0 J) = 0$$

Therefore by Problem 20, we conclude that X_0 is diagonal. Thus for every symmetric matrix H $\exists U_0 \in O(n)$ such that $U_0^{-1}HU_0$ is diagonal.