Note on a Spectral Theorem by Forster

Let $\mathscr{X} = \{\boldsymbol{x} \mid x \in X\} \subseteq \Re^k$ be a collection of unit vectors in $\Re^k$ indexed by a set X of cardinality at least $2k$. We assume that the vectors in $\mathscr{X}$ are in general position, that is, every k-subset of $\mathscr{X}$ is linearly independent. For $A \in \Re^{k\times k}$, we define the following matrix

$$M(A) \equiv \sum_{x\in X} \tfrac{1}{\|A\boldsymbol{x}\|^2}(A\boldsymbol{x} \otimes A\boldsymbol{x}),$$

where $(A\boldsymbol{x} \otimes A\boldsymbol{x})$ denotes the self-adjoint linear operator defined by $(A\boldsymbol{x} \otimes A\boldsymbol{x})(\boldsymbol{y}) = \langle A\boldsymbol{x}, \boldsymbol{y}\rangle A\boldsymbol{x}$. We will show that there exits a matrix $A \in \Re^{k\times k}$ such that

$$\lambda_{\min}(M(A)) = \tfrac{|X|}{k}, \tag{1}$$

where $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a symmetric matrix. Note that any matrix A that satisfies (1) is necessarily invertible, for otherwise the matrix $M(A)$ has smallest eigenvalue 0. Also note that the matrix $M(A)$ has trace $|X|$, and hence the equation (1) implies that all eigenvalues are equal to $|X|/k$.

In the following claim, we gather a few facts about matrices of the form $M(A)$.

Claim 1. *Let $A \in \Re^{k\times k}$. Then,*

1. *$M(A)$ is symmetric and positive semidefinite,*
2. $\operatorname{Tr} M(A) = |X|$,
3. *$\lambda_{\min}(M(A)) = \lambda_{\max}(M(A))$ if and only if $\lambda_{\min}(M(A)) = \frac{|X|}{k}$ (or $\lambda_{\max}(M(A)) = \frac{|X|}{k}$),*
4. *$M(\alpha A) = M(A)$ for any non-zero scalar $\alpha \in \Re$,*
5. *$A \in \mathsf{GL}(k)$ if and only if $M(A) \in \mathsf{GL}(k)$,*

Proof. Items 1–4 can be verified easily. 5.) The range of $M(A)$ is equal to the span of the set $\{A\boldsymbol{x} \mid x \in X\}$. By the general position assumption for $\mathscr{X}$, the span of the set $\{A\boldsymbol{x} \mid x \in X\}$ is equal to $\Re^k$ if A is non-singular. On the other hand, if A is singular, then $\{A\boldsymbol{x} \mid x \in X\}$ cannot span $\Re^k$. □

Lemma 2. *For every non-singular matrix $B \in \Re^{k\times k}$, there exists $\delta > 0$ such that*

$$\lambda_{\min}(M(AB)) \geq \min\left\{\tfrac{|X|}{k}, \lambda_{\min}(M(B)) + \delta\right\},$$

where $A = M(B)^{-1/2}$.

Proof. For $x \in X$, let $\boldsymbol{x}' = \frac{1}{\|B\boldsymbol{x}\|}B\boldsymbol{x}$ be the unit vector in direction $B\boldsymbol{x}$. Note that $(\sum_{x\in X}(\boldsymbol{x}' \otimes \boldsymbol{x}')) = M(B) = A^{-2}$. We may assume $\lambda_{\min} \equiv \lambda_{\min}\left(\sum_{x\in X}(\boldsymbol{x}' \otimes \boldsymbol{x}')\right) < |X|/k$, for otherwise the lemma is trivially true. Also note that $\lambda_{\min} = 1/\lambda_{\max}(A^2) = 1/\lambda_{\max}(A)^2$.

In order to prove the lemma, it remains to show that the matrix $M(AB)-\lambda_{\min} I$ has only positive eigenvalues. Since A is symmetric, we have $(A\boldsymbol{x}'\otimes A\boldsymbol{x}') = A(\boldsymbol{x}'\otimes\boldsymbol{x}')A$ and thus $\sum_{x\in X}(A\boldsymbol{x}'\otimes A\boldsymbol{x}') = A(\sum_{x\in X}(\boldsymbol{x}'\otimes\boldsymbol{x}'))A = I$. Hence,

$$\begin{aligned}
M(A)-\lambda_{\min} I &= \textstyle\sum_{x\in X}\frac{1}{\|A\boldsymbol{x}'\|^2}(A\boldsymbol{x}'\otimes A\boldsymbol{x}') - \lambda_{\min} I \\
&= \textstyle\sum_{x\in X}(\frac{1}{\|A\boldsymbol{x}'\|^2} - \lambda_{\min})(A\boldsymbol{x}'\otimes A\boldsymbol{x}') && (\text{using } \textstyle\sum(A\boldsymbol{x}'\otimes A\boldsymbol{x}') = I) \\
&= \textstyle\sum_{x\in X}\alpha_x(A\boldsymbol{x}'\otimes A\boldsymbol{x}') && (\alpha_x \equiv \tfrac{1}{\|A\boldsymbol{x}'\|^2} - \lambda_{\min}) \\
&\succeq 0. && (\text{using } \alpha_x \geq 0, \text{ because } \|A\boldsymbol{x}\|^2 \leq \lambda_{\max}(A^2) = \lambda_{\min}^{-1})
\end{aligned}$$

Let X_0 denote the set of indices x such that $\alpha_x = 0$. We claim that X_0 has cardinality at most k. Assuming this claim, we can finish the proof of the lemma as follows. If $|X_0| < k$, then there are at least $|X|-k \geq k$ indices such that $\alpha_x > 0$. By the general position assumption, the corresponding set of vectors $\{A\boldsymbol{x}' \mid x \in X\setminus X_0\}$ spans $\Re^k$. Hence for every unit vector $\boldsymbol{y}$, there exists an index $x_1 \in X\setminus X_0$ such that $\langle A\boldsymbol{x}_1', \boldsymbol{y}\rangle \neq 0$. Thus

$$\begin{aligned}
\langle \big(\textstyle\sum_{x\in X}\alpha_x(A\boldsymbol{x}'\otimes A\boldsymbol{x}')\big)\boldsymbol{y},\boldsymbol{y}\rangle &\geq \langle \alpha_{x_1}(A\boldsymbol{x}_1'\otimes A\boldsymbol{x}_1')\boldsymbol{y},\boldsymbol{y}\rangle \\
&= \alpha_{x_1}\langle A\boldsymbol{x}_1',\boldsymbol{y}\rangle^2 > 0.
\end{aligned}$$

It follows that the matrix

$$\textstyle\sum_{x\in X}\alpha_x(A\boldsymbol{x}'\otimes A\boldsymbol{x}') = \sum_{x\in X}\frac{1}{\|A\boldsymbol{x}'\|^2}(A\boldsymbol{x}'\otimes A\boldsymbol{x}') - \lambda_{\min} I = M(AB)-\lambda_{\min} I$$

has only positive eigenvalues, which proves the lemma.

It remains to prove the claim that $|X_0| < k$. For the sake of a contradiction, assume $|X_0| \geq k$. Then there are k linearly independent vectors $\boldsymbol{x}'$ such that $\|A\boldsymbol{x}'\| = \lambda_{\max}(A)$. Thus the eigenspace of A corresponding to $\lambda_{\max}(A)$ has dimension k. It follows that the eigenspace of $A^{-2} = \sum_{x\in X}(\boldsymbol{x}'\otimes\boldsymbol{x}')$ corresponding to $\lambda_{\min} = 1/\lambda_{\max}(A)^2$ has dimension k. Hence $|X| = \mathrm{Tr}\,(A^{-2}) = k\lambda_{\min}$, which contradicts our assumption $\lambda_{\min} < |X|/k$. □

Lemma 3. *Let $\mathscr{A} = \{A^{(\ell)} \mid \ell\in\mathbb{N}\} \subseteq \Re^{k\times k}$ be any sequence of non-singular matrices with $\|A^{(\ell)}\| = 1$. Suppose $\mathscr{A}$ has a subsequence that converges to a singular matrix A. Then, for every $\epsilon > 0$, there exists an $\ell\in\mathbb{N}$ such that*

$$\lambda_{\min}\big(M(A^{(\ell)})\big) < 1+\epsilon.$$

Proof. Suppose that the kernel of A has dimension $d > 0$. Then there exist d ortho-normal vectors $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_d$ such that $\langle A\boldsymbol{x},\boldsymbol{e}_i\rangle = 0$ for every $i\in[d]$ and $x\in X$. Let X_0 denote the set of indices x such that $A\boldsymbol{x} = 0$. Since $\|A\| = \lim_{\ell\to\infty}\|A^{(\ell)}\| = 1$, the matrix A cannot be 0 and hence $d < k$. Therefore, using the general positive assumption, X_0 has cardinality at most d. In the following, we restrict $\mathscr{A}$ to the subsequence that converges to A. Then,

$$\begin{aligned}
&\limsup_{\ell\to\infty} \sum_{i\in[d]} \langle M(A^{(\ell)})\,\boldsymbol{e}_i, \boldsymbol{e}_i\rangle \\
&= \limsup_{\ell\to\infty} \sum_{i\in[d]} \langle \Big(\textstyle\sum_{x\in X} \frac{1}{\|A^{(\ell)}\boldsymbol{x}\|^2} (A^{(\ell)}\boldsymbol{x} \otimes A^{(\ell)}\boldsymbol{x})\Big)\boldsymbol{e}_i, \boldsymbol{e}_i\rangle \\
&= \limsup_{\ell\to\infty} \sum_{i\in[d]} \sum_{x\in X} \tfrac{1}{\|A^{(\ell)}\boldsymbol{x}\|^2} \langle A^{(\ell)}\boldsymbol{x}, \boldsymbol{e}_i\rangle^2 \\
&= \limsup_{\ell\to\infty} \sum_{x\in X_0} \sum_{i\in[d]} \langle \tfrac{1}{\|A^{(\ell)}\boldsymbol{x}\|} A^{(\ell)}\boldsymbol{x}, \boldsymbol{e}_i\rangle^2 && \text{(using } \lim_{\ell\to\infty} \Big\langle \tfrac{1}{\|A^{(\ell)}\boldsymbol{x}\|} A^{(\ell)}\boldsymbol{x}, \boldsymbol{e}_i \Big\rangle = 0 \text{ for } x \notin X_0) \\
&\le \limsup_{\ell\to\infty} \sum_{x\in X_0} 1 && \text{(using } \textstyle\sum_{i\in[d]} \langle \boldsymbol{y}, \boldsymbol{e}_i\rangle^2 \le 1 \text{ for any unit vector } \boldsymbol{y}) \\
&\le d.
\end{aligned}$$

It follows that for every $\epsilon > 0$, there exists $i \in [d]$ and $\ell \in \mathbb{N}$ such that

$$\langle M(A^{(\ell)})\,\boldsymbol{e}_i, \boldsymbol{e}_i\rangle < 1 + \epsilon,$$

which proves the lemma. □

Proof of the Theorem

We will need the following claim.

Claim 4. *At every matrix $A_0 \in \mathsf{GL}(k)$, the following functions are continuous:*

$$g(A) = \lambda_{\min}(M(A)), \qquad f(A) = \lambda_{\min}(M(M(A)^{-1/2}A)).$$

Proof. Follows from the fact that the composition of continuos mappings is continuos. □

Theorem 5. *There exists a matrix $A^* \in \Re^{k\times k}$ such that*

$$\lambda_{\min}\big(M(A^*)\big) = \tfrac{|X|}{k}.$$

Proof. We define a sequence $\mathscr{A} = \{A^{(\ell)} \mid \ell \in \mathbb{N}\} \subseteq \Re^{k\times k}$ of non-singular matrices with $\|A\| = 1$ by the following recurrence

$$A^{(\ell+1)} = \tfrac{1}{\|M(A^{(\ell)})^{-1/2}A^{(\ell)}\|} M(A^{(\ell)})^{-1/2}A^{(\ell)}, \tag{2}$$

where we choose $A^{(1)}$ to be the linear operator that maps the first k vectors in $\mathscr{X}$ to the canonical (orthogonal) basis of $\Re^k$. Note that $M(A^{(\ell)}) = M(M(A^{(\ell-1)})^{-1/2}A^{(\ell-1)})$. Hence, by Lemma 2, the sequence $\{\lambda_{\min}^{(\ell)} \mid \ell \in \mathbb{N}\}$ defined by

$$\lambda_{\min}^{(\ell)} \equiv g(A^{(\ell)}) = \lambda_{\min}\big(M(A^{(\ell)})\big)$$

is strictly increasing in ℓ until it possibly reaches $|X|/k$. Furthermore, we have $\lambda_{\min}^{(1)} \geq 1$, because

$$M(A^{(1)}) = \textstyle\sum_{x \in X} \frac{1}{\|A^{(1)}\boldsymbol{x}\|^2}(A^{(1)}\boldsymbol{x} \otimes A^{(1)}\boldsymbol{x}) \geq \displaystyle\sum_{x \in \{x^1,\ldots,x^k\}} \frac{1}{\|A^{(1)}\boldsymbol{x}\|^2}(A^{(1)}\boldsymbol{x} \otimes A^{(1)}\boldsymbol{x})$$
$$= \sum_{i=1}^{k} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i) = I,$$

where $x^1, \ldots, x^k$ are the first k indices of X, and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ is the canonical basis of $\Re^k$. It follows that $\lambda_{\min}^{(2)} \geq 1 + \epsilon$ for some $\epsilon > 0$.

Let $\mathscr{A}' = \{A^{(\ell(t))} \mid t \in \mathbb{N}\}$ denote a converging subsequence of $\mathscr{A}$. Note that $\mathscr{A}$ has a converging subsequence, because it is contained in the bounded set $\{A \in \Re^{k \times k} \mid \|A\| \leq 1\}$. By Lemma 3 and the observation $\lambda_{\min}^{(\ell)} \geq 1 + \epsilon$ for $\ell > 1$, the limit of $\mathscr{A}'$ is a non-singular matrix A^*. By the continuity of the function f at non-singular matrices,

$$\begin{aligned}
f(A^*) &= \lim_{t\to\infty} f(A^{(\ell(t))}) && \text{(using continuity of } f \text{ at } A^*)\\
&= \lim_{t\to\infty} \lambda_{\min}(M(M(A^{(\ell(t))})^{-1/2} A^{(\ell(t))}))\\
&= \lim_{t\to\infty} \lambda_{\min}\left(M(A^{(\ell(t)+1)})\right) && \text{(using } M(M(A^{(\ell)})^{-1/2}A^{(\ell)}) = M(A^{(\ell+1)}) \text{ for } \ell \in \mathbb{N})\\
&= \lim_{\ell\to\infty} \lambda_{\min}^{(\ell)} && \text{(using convergence of } \{\lambda_{\min}^{(\ell)} \mid \ell \in \mathbb{N}\})\\
&= \lim_{t\to\infty} \lambda_{\min}^{(\ell(t))}\\
&= \lim_{t\to\infty} g(A^{(\ell(t))})\\
&= g(A^*) && \text{(using continuity of } g \text{ at } A^*)
\end{aligned}$$

By Lemma 2, the condition $f(A^*) = g(A^*)$ implies that $\lambda_{\min}(M(A^*)) = \frac{|X|}{k}$, which proves the theorem. □