Proof of Schur's Theorem

David H. Wagner
dhwagnertx@mac.com

In this note, I provide more detail for the proof of Schur's Theorem found in Strang's *Introduction to Linear Algebra*[1]

Theorem 1. *If $\boldsymbol{A}$ is a square real matrix with real eigenvalues, then there is an orthogonal matrix $\boldsymbol{Q}$ and an upper triangular matrix $\boldsymbol{T}$ such that $\boldsymbol{A} = \boldsymbol{QTQ}^{\mathrm{T}}$.*

Proof. Note that $\boldsymbol{A} = \boldsymbol{QTQ}^{\mathrm{T}} \Leftrightarrow \boldsymbol{AQ} = \boldsymbol{QT}$. Let $\boldsymbol{q}_1$ be an eigenvector of norm 1, with eigenvalue λ_1. Let $\boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$ be *any* orthonormal vectors orthogonal to $\boldsymbol{q}_1$. Let $\boldsymbol{Q}_1 = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n]$. Then $\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{Q}_1 = \boldsymbol{I}$, and

$$\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}_1 = \begin{pmatrix} \lambda_1 & \cdots \\ \underline{0} & \boldsymbol{A}_2 \end{pmatrix} \tag{1}$$

Now I claim that $\boldsymbol{A}_2$ has eigenvalues $\lambda_2, \ldots, \lambda_n$. This is true because

$$\begin{aligned} \det(\boldsymbol{A} - \lambda\boldsymbol{I}) &= \det\boldsymbol{Q}_1^{\mathrm{T}} \det(\boldsymbol{A} - \lambda\boldsymbol{I}) \det\boldsymbol{Q}_1 = \det(\boldsymbol{Q}_1^{\mathrm{T}}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{Q}_1) \\ &= \det(\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}_1 - \lambda\boldsymbol{Q}_1^{\mathrm{T}}\boldsymbol{Q}_1) = \det\begin{pmatrix} (\lambda_1 - \lambda) & \cdots \\ \boldsymbol{0} & (\boldsymbol{A}_2 - \lambda\boldsymbol{I}) \end{pmatrix} \\ &= (\lambda_1 - \lambda)\det(\boldsymbol{A}_2 - \lambda\boldsymbol{I}). \end{aligned} \tag{2}$$

So $\boldsymbol{A}_2$ has real eigenvalues, namely $\lambda_2, \ldots, \lambda_n$. Now we proceed by induction. Suppose we have proved the theorem for $n = k$. Then we use this fact to prove the theorem is true for $n = k + 1$. Note that the theorem is trivial if $n = 1$.

So for $n = k+1$, we proceed as above and then apply the known theorem to $\boldsymbol{A}_2$, which is $k \times k$. We find that $\boldsymbol{A}_2 = \boldsymbol{Q}_2\boldsymbol{T}_2\boldsymbol{Q}_2^{\mathrm{T}}$. Now this is the hard part. Let $\boldsymbol{Q}_1$ and $\boldsymbol{A}_2$ be as above, and let

$$\boldsymbol{Q} = \boldsymbol{Q}_1 \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_2 \end{pmatrix} \tag{3}$$

Then

$$\begin{aligned} \boldsymbol{AQ} =& \boldsymbol{AQ}_1 \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_2 \end{pmatrix} = \boldsymbol{Q}_1 \begin{pmatrix} \lambda_1 & \cdots \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_2 \end{pmatrix} \\ &= \boldsymbol{Q}_1 \begin{pmatrix} \lambda_1 & \cdots \\ \boldsymbol{0} & \boldsymbol{A}_2\boldsymbol{Q}_2 \end{pmatrix} = \boldsymbol{Q}_1 \begin{pmatrix} \lambda_1 & \cdots \\ \boldsymbol{0} & \boldsymbol{Q}_2\boldsymbol{T}_2 \end{pmatrix} \\ &= \boldsymbol{Q}_1 \begin{pmatrix} 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots \\ \boldsymbol{0} & \boldsymbol{T}_2 \end{pmatrix} = \boldsymbol{QT}, \end{aligned} \tag{4}$$

where $\boldsymbol{T}$ is upper triangular. So $\boldsymbol{AQ} = \boldsymbol{QT}$, or $\boldsymbol{A} = \boldsymbol{QTQ}^{\mathrm{T}}$. □

That's all, folks!

References

[1] Gilbert Strang, *Introduction to Linear Algebra*, Wellesley-Cambridge Press, Box 812060, Wellesley, Massachusetts 02482, USA, 4th edition, 2009.