PROOF OF SCHUR'S THEOREM

DAVID H. WAGNER

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

Theorem 0.1. *If* $\mathbf{A}$ *is a square real matrix with real eigenvalues, then there is an orthogonal matrix* $\mathbf{Q}$ *and an upper triangular matrix* $\mathbf{T}$ *such that* $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^T$.

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

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

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

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

So $\mathbf{A}_2$ 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 $\mathbf{A}_2$, which is $k \times k$. We find that $\mathbf{A}_2 = \mathbf{Q}_2\mathbf{T}_2\mathbf{Q}_2^T$. Now this is the hard part. Let $\mathbf{Q}_1$ and $\mathbf{A}_2$ be as above, and let

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

Then

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

where $\mathbf{T}$ is upper triangular. So $\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{T}$, or $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^T$. □

That's all, folks!

References

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