Interior eigenvectors of symmetric matrices are saddle points*

G. H. Brown and E. V. Solomonik

Every eigenvector of a symmetric matrix is a critical point of the Rayleigh quotient

$$R\left(\mathbf{A},\mathbf{x}\right)=\frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}},\quad \|\mathbf{x}\|_2^2\neq 0. \tag{1}$$

In fact, this relationship can be used to *define* matrix eigenvalues, with the critical point condition on the Rayleigh quotient being the eigenpair equation $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$. [1]

Obvserve from the eigenpair equation that the eigenvector magnitude is unimportant so long as it is nonzero, which motivates the common choice to 2-normalize eigenvectors. If one does this, all eigenvectors of $\mathbf{A}\in\mathbb{R}^{n\times n}$ lie on the unit $(n-1)$-sphere, and the extreme value theorem can be used to prove that all symmetric matrices with $n\geq 2$ must have at least two eigenvectors, corresponding to the maximum and minimum of the Rayeleigh quotient. However, this only classifies $\mathcal{O}(1)$ critical points. What can be said of the other $n-2$ critical points compirising the interior of the spectrum?

To answer this question, consider an interior eigenpair $(\mathbf{x}_i,\lambda_i)$ of a symmetric matrix $\mathbf{A}$ with $n>2$ and some scaled vector $\beta\mathbf{x}_j$ in the direction of eigenvector $\mathbf{x}_j$. Let the respective norms of $\mathbf{x}_i,\mathbf{x}_j$ be α_i,α_j. [2] The Rayleigh quotient at $\mathbf{x}_i+\beta\mathbf{x}_j$ is

$$\begin{aligned}
R\left(\mathbf{A},\mathbf{x}_i+\beta\mathbf{v}\right)&=\frac{\left(\mathbf{x}_i+\beta\mathbf{x}_j\right)^T\mathbf{A}\left(\mathbf{x}_i+\beta\mathbf{x}_j\right)}{\left(\mathbf{x}_i+\beta\mathbf{x}_j\right)^T\left(\mathbf{x}_i+\beta\mathbf{x}_j\right)} &(2)\\
&=\frac{\mathbf{x}_i^T\mathbf{A}\mathbf{x}_i+2\beta\mathbf{x}_j^T\mathbf{A}\mathbf{x}_i+\beta^2\mathbf{x}_j^T\mathbf{A}\mathbf{x}_j}{\mathbf{x}_i^T\mathbf{x}_i+2\beta\mathbf{x}_j^T\mathbf{x}_i+\beta^2\mathbf{x}_j^T\mathbf{x}_j} &(3)\\
&=\frac{\lambda_i\alpha_i^2+\lambda_j\beta^2\alpha_j^2}{\alpha_i^2+\beta^2\alpha_j^2} &(4)
\end{aligned}$$

where we've expanded and used the symmetry of $\mathbf{A}$ between the first two steps. Between the last two steps we utilized $\mathbf{x}_i$ and $\mathbf{x}_j$ being eigenvectors of $\mathbf{A}$ and the fact that eigenvectors of symmetric matrices are mutually orthogonal. The change in Rayleigh quotient from the original critical point $\mathbf{x}_i$ is then

$$R\left(\mathbf{A},\mathbf{x}_i+\beta\mathbf{x}_j\right)-R(\mathbf{A},\mathbf{x}_i)=\frac{\left(\lambda_j-\lambda_i\right)\beta^2\alpha_j^2}{\alpha_i^2+\beta^2\alpha_j^2} \tag{5}$$

Note that the sign of the quantity above depends only on $\lambda_j-\lambda_i$ since all other quantities are defined to be real numbers. So for an interior eigenpair λ_i there exist at least two unique values of j such that $\lambda_j-\lambda_i<0$ and $\lambda_j-\lambda_i>0$, for concreteness one can choose $j=1,n$ for spectrum $\lambda_1<...<\lambda_n$. [3] Therefore, any interior eigenvector $\mathbf{x}_i$ has an arbitrarily close point (we placed no magnitude restrictions on β) that is larger in Rayleigh quotient and another point that is smaller. This condition defines a saddle point.

[1] The gradient of the Rayleigh quotient is $2\frac{\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}-2\frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\left(\mathbf{x}^T\mathbf{x}\right)^2}\mathbf{x}$. Critical points are defined by a zero gradient, and using the nonzero norm condition on $\mathbf{x}$ one finds $\mathbf{A}\mathbf{x}-R(\mathbf{A},\mathbf{x})\mathbf{x}=\mathbf{0}$. Recognizing that the Rayleigh quotient is a scalar (call it λ), we recover the familiar $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$. At first, it may appear replacing the Rayleigh quotient by some arbitrary scalar could define different conditions if there exists λ such that $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$, but $\lambda\neq R(\mathbf{A},\mathbf{x})$. However, this is not possible, which one can prove by taking the inner product of the of the eigenpair equation with eigenvector $\mathbf{x}$ and rearranging to show that any scalar satisfying the eigenpair equation is precisely the Rayleigh quotient defined by the matrix and eigenvector.

[2] Although employing the extreme value theorem requires a compact domain like the $(n-1)$-sphere, one can classify interior eigenpairs without such a closed domain. Furthermore, it is trivial to reformulate this proof to work on the $(n-1)$-sphere.

[3]* This enforces algebraic multiplicity one for all all eigenvalues. When the spectrum has algebraic multiplicity greater than one at the edges (i.e. $\lambda_1=\lambda_2<...$ for the lower end), it is possible to show that this cluster of eigenpairs are all local minima or maxima by observing that the eigenvectors form a basis for $\mathbb{R}^n$, and so there exist no directions that decrease the Rayleigh quotient, respectively. Similar arguments apply for degenerate maxima at the top of the spectrum.