

# An Extension of the Rayleigh Quotient to the Spectral Radius of Asymmetric Nonnegative Matrices

Lee Altenberg  
Information and Computer Sciences,  
University of Hawai'i at Mānoa, altenber@hawaii.edu

December 29, 2019

## Abstract

The Rayleigh quotient, which provides the classical variational characterization of the spectral radius of Hermitian matrices, can be extended to nonsymmetric nonnegative irreducible matrices,  $\mathbf{A}$ , by the inclusion of a diagonal similarity scaling, to yield the variational formula  $r(\mathbf{A}) = \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x} / (\mathbf{x}^\top \mathbf{x})$ , where  $\mathbf{D}_y$  is the diagonal matrix of the vector  $\mathbf{y}$ . Comparison is made to other variational formulae for the spectral radius.

The classical variational characterization of the spectral radius of Hermitian matrices is Rayleigh's formula. Let  $\mathbf{A} \in \mathbb{C}^{n \times n}$  be an  $n \times n$  Hermitian matrix, and let  $r(\mathbf{A}) = \max \{ |\lambda| : \det(\mathbf{A} - \lambda \mathbf{I}) = 0 \}$  be its spectral radius. Rayleigh's formula is the maximum of the *Rayleigh quotient* (also known as the *Rayleigh-Ritz ratio*):

$$r(\mathbf{A}) = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}^*}{\mathbf{x}^\top \mathbf{x}^*}, \quad (1)$$

where  $\mathbf{x}^*$  is the complex conjugate of  $\mathbf{x}$ .

No such formula exists for general complex square matrices  $\mathbf{A} \in \mathbb{C}^{n \times n}$ . We note that in Rayleigh's formula, both left and right sides of the matrix  $\mathbf{A}$  enter equivalently, which we refer to here as *structural symmetry*.

Variational formulae do exist for the class of irreducible nonnegative square matrices. In contrast with the structurally symmetric form of the Rayleigh quotient, the

most well-known formulae for the spectral radius of irreducible nonnegative matrices are *structurally asymmetric*, in that the two sides of the matrix enter in different roles. There is the classical Collatz-Wielandt formula:

$$r(\mathbf{A}) = \max_{\mathbf{x} > \mathbf{0}} \min_i \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} = \min_{\mathbf{x} > \mathbf{0}} \max_i \frac{[\mathbf{A}\mathbf{x}]_i}{x_i},$$

where only one side of the matrix is involved in vector multiplication (although either side could be chosen), and the Donsker-Varadhan-Friedland formula (Donsker and Varadhan, 1975; Friedland, 1981):

$$r(\mathbf{A}) = \sup_{\mathbf{p} \in \mathcal{P}} \inf_{\mathbf{x} > \mathbf{0}} \sum_{i=1}^n p_i \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} = \sup_{\mathbf{p} \in \mathcal{P}} \inf_{\mathbf{x} > \mathbf{0}} (\mathbf{p} \circ \mathbf{x}^{-1})[\mathbf{A}\mathbf{x}]_i, \quad (2)$$

where  $\mathbf{x}^{-1} := [1/x_i]_{i=1}^n$ ,  $\mathbf{p} \circ \mathbf{x}^{-1} = [p_i/x_i]_{i=1}^n$ ,  $\circ$  being the elementwise Hadamard product, and the simplex is represented as  $\mathcal{P} = \{\mathbf{p} : p_i \geq 0, \sum_{i=1}^n p_i = 1\} \subset \mathbb{R}^n$ . Here the asymmetry enters in the different spaces of variation on the left and right vectors multiplying  $\mathbf{A}$ .

Fiedler et al. (1985) introduced a structurally symmetric variational formula for the spectral radius:

$$r(\mathbf{A}) = \max_{\mathbf{z} > \mathbf{0}} \min_{\mathbf{x} > \mathbf{0}, \mathbf{y} > \mathbf{0}, \mathbf{x} \circ \mathbf{y} = \mathbf{z}} \frac{\mathbf{y}^\top \mathbf{A} \mathbf{x}}{\mathbf{y}^\top \mathbf{x}}.$$

Here, both sides of the matrix  $\mathbf{A}$  are multiplied by vector variables. Asymmetries in  $\mathbf{A}$  manifest as inequality between  $\mathbf{y}$  and  $\mathbf{x}$  in the vectors at which the min and max are attained.

A change of variables shows that the Donsker-Varadhan-Friedland formula and the Fiedler formula are actually equivalent. The constraint  $\mathbf{x} \circ \mathbf{y} = \mathbf{z}$  in the Fiedler formula means that the degrees of freedom in  $\{\mathbf{x}, \mathbf{y}\}$  are not actually  $2n$ , but only  $n$ , since there are  $n$  constraints  $y_i = z_i/x_i$ . An additional  $n$  degrees of freedom come from variation of  $\mathbf{z}$ . The normalization by  $\mathbf{x}^\top \mathbf{y}$  brings the total variational degrees of freedom to  $2n - 1$ . With the constraints incorporated, the Fiedler formula becomes

$$r(\mathbf{A}) = \max_{\mathbf{z} > \mathbf{0}} \min_{\mathbf{x} > \mathbf{0}} \frac{(\mathbf{z} \circ \mathbf{x}^{-1})^\top \mathbf{A} \mathbf{x}}{\mathbf{z}^\top \mathbf{e}} = \max_{\mathbf{z} > \mathbf{0}} \min_{\mathbf{x} > \mathbf{0}} \sum_{i=1}^n \frac{z_i}{\mathbf{z}^\top \mathbf{e}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i},$$

which is the Donsker-Varadhan-Friedland formula since  $\mathbf{z}/\mathbf{z}^\top \mathbf{e} \in \mathcal{P}$ , which likewise has  $2n - 1$  variational degrees of freedom. The inf and sup are replaced by min and

max because, as Fiedler et al. (1985) show, the max and min are actually attained for specific  $\mathbf{x} > \mathbf{0}, \mathbf{y} > \mathbf{0}$ .

In this brief note, we shall see that the methods used by Fiedler et al. (1985) to prove Fiedler's formula can be used to prove another structurally symmetric variational formula for the spectral radius. The formula exhibits the most natural way to handle asymmetries in the matrix, which is through a diagonal similarity scaling. As in the other variational formulae, it also has  $2n - 1$  degrees of variational freedom. What is perhaps most remarkable is that the formula is not already widely known.

Let  $\mathbf{y} \in \mathbb{R}^n$  and define  $\mathbf{D}_\mathbf{y} := \mathbf{diag}[\mathbf{y}]$  to be the diagonal matrix with diagonal elements  $y_i$ . A diagonal similarity scaling of a matrix  $\mathbf{A}$  is the product  $\mathbf{D}_\mathbf{y}\mathbf{A}\mathbf{D}_\mathbf{y}^{-1}$  where we assume  $\mathbf{y} > \mathbf{0}$ .

**Theorem 1** (Structurally Symmetric Variational Formula for the Spectral Radius).

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  be a nonnegative irreducible  $n \times n$  matrix, with Perron root  $r(\mathbf{A})$ , and Perron vectors  $\mathbf{u}^\top \mathbf{A} = r(\mathbf{A})\mathbf{u}^\top$  and  $\mathbf{A}\mathbf{v} = r(\mathbf{A})\mathbf{v}$ , normalized so that  $\mathbf{u}^\top \mathbf{v} = 1$ . Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ . Then

$$r(\mathbf{A}) = \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_\mathbf{y} \mathbf{A} \mathbf{D}_\mathbf{y}^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}. \quad (3)$$

Moreover, the sup and inf are attained at  $\hat{\mathbf{x}}^\top \mathbf{D}_{\hat{\mathbf{y}}} = \mathbf{u}^\top$  and  $\mathbf{D}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{x}} = \mathbf{v}$ , i.e.  $\hat{x}_i \hat{y}_i = u_i$  and  $\hat{x}_i / \hat{y}_i = v_i$ , hence  $u_i v_i = \hat{x}_i^2$ , which have solutions

$$\begin{aligned} \hat{x}_i &= \sqrt{u_i v_i}, \\ \hat{y}_i &= u_i / \hat{x}_i = u_i / \sqrt{u_i v_i} = \sqrt{u_i / v_i} \\ &= x_i / v_i = \sqrt{u_i v_i} / v_i = \sqrt{u_i / v_i}, \end{aligned}$$

for which we find that

$$\frac{\hat{\mathbf{x}}^\top \mathbf{D}_{\hat{\mathbf{y}}} \mathbf{A} \mathbf{D}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{x}}}{\hat{\mathbf{x}}^\top \hat{\mathbf{x}}} = \frac{\mathbf{u}^\top \mathbf{A} \mathbf{v}}{\mathbf{u}^\top \mathbf{v}} = \mathbf{u}^\top \mathbf{A} \mathbf{v} = r(\mathbf{A}) \mathbf{u}^\top \mathbf{v} = r(\mathbf{A}).$$

*Proof.* We follow the method of proof of Fiedler's theorem from Fiedler et al. (1985, pp. 83–86). First, we show that for any  $\mathbf{x} > \mathbf{0}$ ,  $\inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_\mathbf{y} \mathbf{A} \mathbf{D}_\mathbf{y}^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq r(\mathbf{A})$ . Given  $\mathbf{x}$ , let  $\mathbf{y} = \mathbf{u}/\mathbf{x}$ . Then

$$\begin{aligned} \frac{\mathbf{x}^\top \mathbf{D}_\mathbf{y} \mathbf{A} \mathbf{D}_\mathbf{y}^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} &= \frac{\mathbf{x}^\top \mathbf{D}_{\mathbf{u}/\mathbf{x}} \mathbf{A} \mathbf{D}_{\mathbf{x}/\mathbf{u}} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \frac{\mathbf{u}^\top \mathbf{A} \mathbf{D}_{\mathbf{x}/\mathbf{u}} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = r(\mathbf{A}) \frac{\mathbf{u}^\top \mathbf{D}_{\mathbf{x}/\mathbf{u}} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \\ &= r(\mathbf{A}) \frac{\mathbf{x}^\top \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = r(\mathbf{A}), \end{aligned}$$

hence

$$\phi(\mathbf{x}) := \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \leq r(\mathbf{A}). \quad (4)$$

Second, we show that  $\sup_{\mathbf{x} > \mathbf{0}} \phi(\mathbf{x}) \geq r(\mathbf{A})$ . Consider  $\phi(\sqrt{\mathbf{u}\mathbf{v}})$ :

$$\begin{aligned} \phi(\sqrt{\mathbf{u}\mathbf{v}}) &= \inf_{\mathbf{y} > \mathbf{0}} \frac{\sqrt{\mathbf{u}\mathbf{v}}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \sqrt{\mathbf{u}\mathbf{v}}}{\sqrt{\mathbf{u}\mathbf{v}}^\top \sqrt{\mathbf{u}\mathbf{v}}} = \inf_{\mathbf{y} > \mathbf{0}} \frac{\sqrt{\mathbf{u}\mathbf{v}}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \sqrt{\mathbf{u}\mathbf{v}}}{\mathbf{u}^\top \mathbf{v}} \\ &= \inf_{\mathbf{y} > \mathbf{0}} \sqrt{\mathbf{u}\mathbf{v}}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \sqrt{\mathbf{u}\mathbf{v}}. \end{aligned}$$

Substitute  $\mathbf{z} = \mathbf{y} \circ \sqrt{\mathbf{v}/\mathbf{u}}$ , so  $\mathbf{y} = \mathbf{z} \circ \sqrt{\mathbf{u}/\mathbf{v}}$ . Then

$$\begin{aligned} \phi(\sqrt{\mathbf{u}\mathbf{v}}) &= \inf_{\mathbf{y} > \mathbf{0}} \sqrt{\mathbf{u}\mathbf{v}}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \sqrt{\mathbf{u}\mathbf{v}} = \inf_{\mathbf{z} > \mathbf{0}} \sqrt{\mathbf{u}\mathbf{v}}^\top \mathbf{D}_{(\mathbf{z} \circ \sqrt{\mathbf{u}/\mathbf{v}})} \mathbf{A} \mathbf{D}_{(\mathbf{z} \circ \sqrt{\mathbf{u}/\mathbf{v}})}^{-1} \sqrt{\mathbf{u}\mathbf{v}} \\ &= \inf_{\mathbf{z} > \mathbf{0}} \mathbf{z}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{z}^{-1}. \end{aligned}$$

Fiedler et al. (1985, p. 84) and Eaves et al. (1985, Corollary 3) showed that matrices diagonally scaled by their Perron vectors,  $\mathbf{D}_u \mathbf{A} \mathbf{D}_v$ , exhibit for all  $\mathbf{z} > \mathbf{0}$ ,

$$\mathbf{z}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{z}^{-1} \geq \mathbf{e}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{e}, \quad (5)$$

hence,  $\inf_{\mathbf{z} > \mathbf{0}} \mathbf{z}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{z}^{-1}$  is attained at  $\mathbf{z} = \mathbf{e}$ , which yields

$$\phi(\sqrt{\mathbf{u}\mathbf{v}}) = \inf_{\mathbf{z} > \mathbf{0}} \mathbf{z}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{z}^{-1} = \mathbf{e}^\top \mathbf{D}_u \mathbf{A} \mathbf{D}_v \mathbf{e} = \mathbf{u}^\top \mathbf{A} \mathbf{v} = r(\mathbf{A}).$$

Therefore,

$$\sup_{\mathbf{x} > \mathbf{0}} \phi(\mathbf{x}) \geq \phi(\sqrt{\mathbf{u}\mathbf{v}}) = r(\mathbf{A}). \quad (6)$$

Since by (4),  $\sup_{\mathbf{x} > \mathbf{0}} \phi(\mathbf{x}) \geq r(\mathbf{A})$  and by (6)  $\phi(\mathbf{x}) \leq r(\mathbf{A})$  for all  $\mathbf{x} > \mathbf{0}$ , we obtain

$$\sup_{\mathbf{x} > \mathbf{0}} \phi(\mathbf{x}) = \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = r(\mathbf{A}).$$

□

**Remark:** An interpretation of this result is that the diagonal similarity scaling  $\mathbf{D}_{\sqrt{\mathbf{u}/\mathbf{v}}} \mathbf{A} \mathbf{D}_{\sqrt{\mathbf{v}/\mathbf{u}}}$  renders  $\mathbf{A}$  “as symmetric” as possible, to the point where it is fit to be used in Rayleigh’s quotient. The vector  $\sqrt{\mathbf{u}/\mathbf{v}}$  could be thought of as a measure of the asymmetry of  $\mathbf{A}$ .

This result is readily extended to the spectral bound  $s(\mathbf{A}) := \max \{ \operatorname{Re}(\lambda) : \det(\mathbf{A} - \lambda \mathbf{I}) = 0 \}$ , of essentially nonnegative matrices, in which off-diagonal elements are nonnegative and diagonal elements are real.

**Corollary 1.** Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  be an irreducible essentially nonnegative  $n \times n$  matrix. Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ . Then the spectral bound of  $\mathbf{A}$  is

$$s(\mathbf{A}) = \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}.$$

*Proof.* Let  $t = \min \{A_{ii}\}$ . Then  $\mathbf{B} = \mathbf{A} - t\mathbf{I} \geq \mathbf{0}$ , and  $r(\mathbf{B}) = s(\mathbf{A}) - t$ . Hence

$$\begin{aligned} s(\mathbf{A}) &= r(\mathbf{B}) + t = \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y (\mathbf{A} - t\mathbf{I}) \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} + t \\ &= \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \left( \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} - t \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{I} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \right) + t \\ &= \sup_{\mathbf{x} > \mathbf{0}} \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} - t + t. \end{aligned}$$

□

Rayleigh's formula is recovered from (3) in the case of symmetric nonnegative irreducible matrices,  $\mathbf{S}$ . We have

$$\phi(\mathbf{x}) = \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{D}_y \mathbf{S} \mathbf{D}_y^{-1} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{y}^\top \mathbf{D}_x \mathbf{S} \mathbf{D}_x \mathbf{y}^{-1}}{\mathbf{x}^\top \mathbf{x}}$$

Since  $\mathbf{D}_x \mathbf{S} \mathbf{D}_x$  is symmetric, it is *line-sum symmetric*, i.e.  $\mathbf{D}_x \mathbf{S} \mathbf{D}_x \mathbf{e} = (\mathbf{D}_x \mathbf{S} \mathbf{D}_x)^\top \mathbf{e}$ . Hence by (5),  $\inf_{\mathbf{y} > \mathbf{0}} \mathbf{e}^\top \mathbf{D}_y (\mathbf{D}_x \mathbf{S} \mathbf{D}_x) \mathbf{D}_y^{-1} \mathbf{e}$  is attained for  $\mathbf{y} = \mathbf{e}$ . Thus

$$\phi(\mathbf{x}) = \inf_{\mathbf{y} > \mathbf{0}} \frac{\mathbf{y}^\top \mathbf{D}_x \mathbf{S} \mathbf{D}_x \mathbf{y}^{-1}}{\mathbf{x}^\top \mathbf{x}} = \frac{\mathbf{e}^\top \mathbf{D}_x \mathbf{S} \mathbf{D}_x \mathbf{e}}{\mathbf{x}^\top \mathbf{x}} = \frac{\mathbf{x}^\top \mathbf{S} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

so

$$r(\mathbf{S}) = \sup_{\mathbf{x} > \mathbf{0}} \phi(\mathbf{x}) = \sup_{\mathbf{x} > \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{S} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^\top \mathbf{S} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}},$$

which is Rayleigh's formula, where  $\mathbf{x} > \mathbf{0}$  can be weakened to  $\mathbf{x} \neq \mathbf{0}$  since by Perron-Frobenius theory the maximum is attained for  $\mathbf{x} > \mathbf{0}$ .

While the Rayleigh quotient can be used to obtain all of the eigenvalues of Hermitian matrices, as is done in the Courant-Fischer theorem (Horn and Johnson, 1985, p. 179), no such direct use of (3) for nonsymmetric matrices is possible, because inf and sup do not apply to the complex eigenvalues that may exist.

**Conclusion:** The formula (3) in Theorem 1 wraps the matrix  $\mathbf{A}$  in a diagonal similarity scaling  $\mathbf{D}_y \mathbf{A} \mathbf{D}_y^{-1}$ , which is able to accommodate any asymmetries in  $\mathbf{A}$ ,

and when the rescaled product is placed into the Rayleigh quotient, its maximization produces the spectral radius. The formula applies to irreducible nonnegative matrices, but much of Perron-Frobenius theory has been extended beyond this class of matrices. It remains to be addressed just how widely (3) remains true for these extended matrix classes, and how it may be applied to resolvent positive operators more generally.

## Acknowledgments

The author thanks Marcus W. Feldman for support from the Stanford Center for Computational, Evolutionary and Human Genomics and the Morrison Institute for Population and Resources Studies, Stanford University, and the Mathematical Biosciences Institute at The Ohio State University, for its support through National Science Foundation Award #DMS 0931642. This work was motivated by pursuit of a conjecture by Joel E. Cohen on Levinger's Theorem.

## References

- Donsker, M. D. and Varadhan, S. R. S. 1975. On a variational formula for the principal eigenvalue for operators with maximum principle. *Proceedings of the National Academy of Sciences U.S.A.*, 72:780–783.
- Eaves, B., Hoffman, A., Rothblum, U., and Schneider, H. 1985. Line-sum-symmetric scalings of square nonnegative matrices. *Mathematical Programming Study*, 25(Essays in Honor of George B. Dantzig Part II):124–141.
- Fiedler, M., Johnson, C. R., Markham, T. L., and Neumann, M. 1985. A trace inequality for M-matrices and the symmetrizability of a real matrix by a positive diagonal matrix. *Linear Algebra and Its Applications*, 71:81–94.
- Friedland, S. 1981. Convex spectral functions. *Linear and Multilinear Algebra*, 9:299–316.
- Horn, R. A. and Johnson, C. R. 1985. *Matrix Analysis*. Cambridge University Press, Cambridge.