# ON THE RAYLEIGH-RITZ QUOTIENT

Alina PETRESCU-NIŢĂ[1]

*In linear algebra, the Rayleigh-Ritz quotient is defined for any operator of a Hilbert space. In the case of selfadjoint operators, it allows variational characterizations of their eigenvalues. In the present paper, properties are presented for this relationship in the case of general operators, including the original result which shows that the Rayleigh-Ritz quotientis give an estimate of the "variation speed" for the operator.*

**Keywords:** self-adjoinct operator, eigenvalues, Hermitian operator, pre-Hilbert space

**MSC2000:** Primary 47B15, 47B25. Secondary 66N25, 65N30

## 1. Introduction

Let us assume $(E, \langle\cdot,\cdot\rangle)$ as a complex $n$-dimensional Euclidean space. Therefore, $E$ is a vector space over $\mathbb{C}$, equipped with a scalar product. A linear operator $f : E \to E$ is defined as *orthogonal* when it preserves the scalar product, that is $\forall x, y \in E$, $\langle f(x), f(y)\rangle = \langle x, y\rangle$, and *selfadjoint* ($\equiv$ symmetrical) when $\forall x, y \in E$, $\langle f(x), y\rangle = \langle x, f(y)\rangle$. We define, as usual, $f^*$ the adjoint of $f$ (its defining property being $\forall x, y \in E$, $\langle f^*(x), y\rangle = \langle x, f(y)\rangle$), and we show that $f$ is orthogonal if and only if $f$ is a linear isomorphism and $f^{-1} = f^*$; hence $f$ is selfadjoint if and only if $f^* = f$.

When we use matrices, take $\mathcal{B} = \{e_1, e_2, \ldots, e_n\}$ an orthonormal basis of $E$ and consider the square matrix $A = M_f^{\mathcal{B}} = (f(e_1)|f(e_2)|\ldots|f(e_n))$ of the operator $f$ in relation to $\mathcal{B}$, then the matrix of $f^*$ in relation to $\mathcal{B}$ is of $\bar{A}^T$ (denoted as $A^*$). It should be reminded that a square matrix $A \in M_n(\mathbb{C})$ is called *orthogonal* when $A^* \cdot A = \dot{I}_n$, that is when $A$ is invertible and $A^{-1} = A^*$, and as a *Hermitian* matrix when $A = A^*$.

Then, the operator $f : E \to E$ is orthogonal if and only if the matrix associated to $f$ is orthogonal in any orthonormal basis; similarly, $f$ is selfadjoint if and only if, the matrix of $f$ is a Hermitian in any orthonormal basis. It is known that the eigenvalues of Hermitian operators are real.

We recall that for any linear operator $f : E \to E$ of a complex Euclidean space and for any non-zero vector $x \in E$, the *Rayleigh-Ritz quotient* of $f$ in $x$ is defined as

[1]Lecturer, Department of Mathematical Methods and Models, Faculty of Applied Sciences, University "POLITEHNICA" of Bucharest, Splaiul Independenţei 313,, 060042, Bucharest, Romania, E-mail: alina.petrescu@upb.ro

$$R_f(x) = \frac{\langle f(x), x\rangle}{||x||^2}; \quad [1],[2],[4],[5],[6] \tag{1}$$

When $\mathcal{B}$ is an orthonormal basis of $E$ and $A = M_f^{\mathcal{B}}$ is the matrix associated to $f$ in relation to $\mathcal{B}$, then for any vector $x \in E$ (identified with the column-vector $X$ of the coordinates of $x$ in the basis $\mathcal{B}$) the relationship $f(x) = A \cdot X$ takes place. We will define $X^* = \bar{X}^T$ the transposed of the conjugate of the vector $X$. In addition, for every $x, y \in E$, we have $\langle x, y\rangle = X^* \cdot Y$. Hence, $||x||^2 = X^* \cdot X$. Considering these identifications, the Rayleigh-Ritz quotient may be written equivalently as follows:

$$R_f(x) = \frac{\langle A \cdot X, X\rangle}{||x||^2} = \frac{(A \cdot X)^* \cdot X}{X^* \cdot X} = \frac{X^* \cdot A^* \cdot X}{X^* \cdot X}, \quad \text{for any } x \neq 0. \tag{2}$$

**Example 1.1.** 1) Let us assume that $E = \mathbb{R}^2$ and $\rho_\theta : E \to E$ is the rotation operator with an angle $\theta$ around the origin. Then, for every $x \in \mathbb{R}^2 \setminus \{0\}$, $\rho_\theta(x)$, is the rotated of $x$ with an angle $\theta$; relative to the Euclidean scalar product we have $\langle \rho_\theta(x), x\rangle = ||x|| \cdot ||x|| \cdot \cos\theta$ and $R_{\rho_\theta(x)} = \cos\theta$ is independent of $x$.

2) Let us assume that $v \in E$ is a non-zero vector; since it is linearly independent, it may be completed to an orthogonal basis $\{v, u_2, \dots, u_n\}$ of $E$ (by means of the Gram-Schmidt process). Then, the vector $f(v)$ will be written in form of

$$f(v) = a \cdot v + \sum_{k=2}^{n} b_k \cdot u_k \quad (\text{with } a, b_k \in \mathbb{C}),$$

hence

$$\langle f(v), v\rangle = \left\langle a \cdot v + \sum_{k=2}^{n} b_k \cdot u_k, v \right\rangle = a\langle v, v\rangle + \sum_{k=2}^{n} b_k \cdot 0 = a \cdot v^* \cdot v = a \cdot ||v||^2.$$

In this case, according to (1), $R_f(v) = \dfrac{\langle f(v), v\rangle}{||v||^2} = a$, and it results that the Rayleigh-Ritz quotient is precisely the scalar projection of $f(v)$ along $v$.

One classic result consists in the following

**Theorem 1.1.** Theorem (Rayleigh-Ritz [2]). *When $A \in M_n(\mathbb{C})$ is a Hermitian matrix and its eigenvalues are arranged in increasing order: $\lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n = \lambda_{\max}$, then for any possible vector $X \in \mathbb{C}^n$, the following take place:*

$$\lambda_1 X^* \cdot X \leq X^* \cdot A \cdot X \leq \lambda_n X^* \cdot A \cdot X, \tag{3}$$

$$\begin{aligned} \lambda_{\min} &= \min_{X \neq 0} \frac{X^* \cdot A \cdot X}{X^* \cdot X} = \min_{||X||=1} X^* \cdot A \cdot X; \\ \lambda_{\max} &= \max_{X \neq 0} \frac{X^* \cdot A \cdot X}{X^* \cdot X} = \max_{||X||=1} X^* \cdot A \cdot X. \end{aligned} \tag{4}$$

This theorem happens for every selfadjoint operator $f : E \to E$ and it is expressed simply through the proposition (1). Namely as follows:

**Proposition 1.1.** *The function* $R_f : E \setminus \{0\} \to \mathbb{C}$ *has real values and, moreover, for the eigenvalues (the lowest and the highest one) of* $f$*, we have*

$$\lambda_{\min} = \max_{\substack{x \in E \\ x \neq 0}} R_f(x) \quad \text{and} \quad \lambda_{\max} = \max_{\substack{x \in E \\ x \neq 0}} R_f(x) \tag{5}$$

The result can be refined, considering an orthonormal basis $\mathcal{B}$ of $E$ in relation to which the operator $f$ is diagonal. It is known that the matrix $A = M_f^{\mathcal{B}}$ will be decomposed as follows: $A = U \cdot D \cdot U^*$, where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and $U = (u_1|u_2|\ldots|u_n)$ is a matrix having in its column the orthonormal eigenvectors of $A$. Considering an orthogonal vector $x$ on $u_1$, we have the following results:

$$x^* \cdot A \cdot x = \sum_{k=1}^{n} \lambda_k |(U^* \cdot x)_k|^2 = \sum_{k=2}^{n} \lambda_k |(u_k^* \cdot x)|^2,$$

and so

$$x^* \cdot A \cdot x \geq \lambda_2 \sum_{k=2}^{n} |u_k^* \cdot x|^2 = \lambda_2 \sum_{k=1}^{n} |(U^* \cdot x)_k|^2 = \lambda_2 \cdot x^* \cdot x.$$

Therefore, $\lambda_2 = \min\limits_{\substack{x \neq 0 \\ x \perp \mu_1}} R_f(x)$.

**Remark 1.1.** If $A \in M_{m,n}(\mathbb{C})$ is a rank-$k$ matrix, then $A^* \cdot A$ is a square, non-negative, definite matrix, (because $X^* \cdot (A^* \cdot A) \cdot X = (A \cdot X)^* \cdot (A \cdot X) = ||A \cdot X||^2 \geq 0$ for any $X$). Then, the eigenvalues of $A^* \cdot A$ are non-negative. The non-zero square roots (as many as $k$) of these eigenvalues are precisely the singular values of the matrix $A$ and they form the diagonal of a matrix $\Sigma$ with the remaining elements null; moreover, the singular decomposition of $A$ takes the form $A = V \cdot \Sigma \cdot W^*$, where $V \in M_m(\mathbb{C})$ and $W \in M_n(\mathbb{C})$ are unitary matrices. Such decomposition allows the direct calculation of the pseudo-inverse of $A$, namely $A^+ = V \cdot \Sigma^+ \cdot W^*$ [3]. The Rayleigh-Ritz quotient for the matrix $A^* \cdot A$ is, according to $(1.1)'$,

$$\frac{X^* \cdot (A^* \cdot A)^* \cdot X}{X^* \cdot X} = \frac{X^* \cdot A^* \cdot A^* \cdot X}{X^* \cdot X} = \frac{(A^* \cdot X)^* \cdot A \cdot X}{X^* \cdot X} =$$

$$= \frac{||A \cdot X||^2}{||X||^2} = \left( \frac{||A \cdot X||}{||X||} \right)^2 .$$

Its maximum is the square of the "augmentation" of the input-output linear system $\mathbb{C}^n \to \mathbb{C}^m$ defined by the matrix $A$ [4].

## 2. The Rayleigh-Ritz quotient for linear operators

Let us assume that $E$ is a complex pre-Hilbert space and $f : E \to E$ a linear, not necessarily selfadjoint operator.

**Proposition 2.1.** *The function*

$$\rho : E \setminus \{0\} \to \mathbb{R}, \quad \rho(x) = |R_f(x)|$$

*is bounded and it attains its bounds.*

*Proof.* Let us define $S = \{x \in E| \; ||x|| = 1\}$, the unit sphere. $E$ is a metric space (isometric with $\mathbb{C}^n \cong \mathbb{R}^{2n}$) and $S$ is a compact set. Moreover, the function $\rho$ is continues on $E \setminus \{0\}$, and therefore on $S \subset E \setminus \{0\}$, too. Consequently, the restriction $\rho|_S$ is bounded and it is touching its bounds $m, M$; $m = \inf\limits_{x \in S} |\rho(x)|$. When $x \in E \setminus \{0\}$ is a vector and $u = \dfrac{x}{||x||}$ is the versor of $x$, then $u \in S$ and $x = ||x||u$, and we have

$$\langle f(x), x\rangle = \langle ||x|| \cdot f(u), ||x|| \cdot u\rangle = ||x||^2 \cdot \langle f(u), u\rangle = ||x||^2 \cdot R_f(u) \cdot ||u||^2 = ||x||^2 \cdot R_f(u).$$

Therefore, $\dfrac{\langle f(x), x\rangle}{||x||^2} = R_f(u)$; hence $R_f(x) = R_f(u)$ and $\rho(x) = \rho(u)$. Since $u \in S$, it results that $\rho(x) \geq m$. Similarly, we can show that for whatever $x \in E \setminus \{0\}$, we have $\rho(x) \leq M$. This means that the function $\rho$ is bounded over $E \setminus \{0\}$, while its margins are attain on $S$. □

**Remark 2.1.** If $f$ is a selfadjoint operator, then the proposition 2.1. is more precise; namely, according to the Rayleigh-Ritz theorem, the function $R_f$ has real values, its extreme global values being on $E \setminus \{0\}$, precisely $\lambda_{\min}$ and $\lambda_{\max}$ (that is the extreme eigenvalue of the operator $f$); [5],[6].

**Proposition 2.2.** *When the operator $f$ is orthogonal, then* $\forall x \in E \setminus \{0\}$, $|R_f(x)| \leq 1$.

*Proof.* We have $\forall x \in E$, $||f(x)||^2 = \langle f(x), f(x)\rangle = \langle x, f^*(f(x))\rangle = ||x||^2$, since $f^* \circ f = 1_E$. Therefore, $||f(x)|| = ||x||$, and according to Schwarz's inequality $|\langle f(x), x\rangle| \leq ||f(x)|| \cdot ||x|| = ||x||^2$, and hence

$$|R_f(x)| = \frac{|\langle f(x), x\rangle|}{||x||^2} \leq 1.$$

□

## 3. The variation speed of the Rayleigh-Ritz quotient

Let us assume that $E$ is a real pre-Hilbert space and $f : E \to E$ a linear operator. If $\mathcal{B}$ is an orthonormal basis of $E$ and $A = M_f^{\mathcal{B}}$, then $\forall x, y \in E$, we have $f(x) = A \cdot X$; $\langle x, y\rangle = X^T \cdot Y$. Then, also, $\langle x, f(y)\rangle = X^T \cdot A \cdot Y$ and $\langle y, f(x)\rangle = Y^T \cdot A \cdot X = X^T \cdot A^T \cdot Y$ (the latter relationship being between real numbers and for $a \in \mathbb{R}$, $a^T = a$). Then

$$\langle x, f(y)\rangle + \langle y, f(x)\rangle = X^T \cdot (A + A^T) \cdot Y. \tag{6}$$

On the other hand, $\langle f(x), x\rangle = (A \cdot X)^T \cdot X = X^T \cdot A^T \cdot X$; and so considering the gradients we have,

$$\text{grad}\,\langle f(x), x\rangle = \text{grad}\,(X^T \cdot A^T \cdot X) = X^T \cdot (A + A^T). \tag{7}$$

Comparing the relationships (6) and (7), we get the following result

$$\langle x, f(y)\rangle + \langle f(x), y\rangle = \langle \text{grad}\,\langle f(x), x\rangle, y\rangle. \tag{8}$$

In case that $f = 1_E$, the relationship (8) becomes

$$\langle \text{grad}\,\langle x, x\rangle, y\rangle = 2\langle x, y\rangle. \tag{9}$$

On the other hand, according to (1), and applying the formulas

$$\text{grad}\,\frac{\varphi}{\psi} = \frac{\psi \text{grad}\varphi - \varphi \text{grad}\psi}{\psi^2}\ (\psi \neq 0) \quad \text{and} \quad \text{grad}\,||x||^2 = 2x,$$

we have

$$\begin{aligned} \text{grad}\,R_f(x) &= \frac{||x||^2 \text{grad}\,\langle f(x), x\rangle - \langle f(x), x\rangle \text{grad}\,||x||^2}{||x||^4} = \\ &= \frac{1}{||x||^2}\text{grad}\,\langle f(x), x\rangle - \frac{2x}{||x||^4}\langle f(x), x\rangle. \end{aligned}$$

Therefore, for every $x, y \in E$; $x \neq 0$, we have

$$\langle \text{grad}\,R_f(x), y\rangle = \frac{1}{||x||^2}\langle \text{grad}\,\langle f(x), x\rangle, y\rangle - \frac{2}{||x||^4}\langle f(x), x\rangle \cdot \langle x, y\rangle.$$

According to (8), we have

$$\langle \text{grad}\,\langle f(x), x\rangle, y\rangle = \langle x, f(y)\rangle + \langle f(x), y\rangle = \langle f^*(x), y\rangle + \langle f(x), y\rangle;$$

So that

$$\begin{aligned} \langle \text{grad}\,R_f(x), y\rangle &= \frac{1}{||x||^2}(\langle f(x) + f^*(x), y\rangle) - \frac{2}{||x||^4}||x||^2 R_f(x) \cdot \langle x, y\rangle = \\ &= \frac{1}{||x||^2}\left[\langle f(x) + f^*(x), y\rangle - 2R_f(x) \cdot \langle x, y\rangle\right] = \\ &= \frac{1}{||x||^2}\left\langle f(x) + f^*(x) - 2R_f(x) \cdot x, y\right\rangle. \end{aligned}$$

Since $y$ is a general random vector in $E$, we have the following result

**Proposition 3.1.** *If $f : E \to E$ is a linear operator of a real pre-Hilbert space, then for any $x \in E$; $x \neq 0$, we have*

$$\text{grad}\,R_f(x) = \frac{1}{||x||^2}(f(x) + f^*(x) - 2R_f(x) \cdot x). \tag{10}$$

When we know the gradient of a scalar field $\varphi$, it is possible to determine the density of $\varphi$ along every direction of a versor $s$, namely $\dfrac{d\varphi}{ds}(x) = \langle \text{grad}\,\varphi(x), s\rangle$; this enables the determination of the rate of change of $\varphi$ along the direction of $s$.

**Corollary 3.1.** *When the linear operator $f : E \to E$ is selfadjoint and $x \in E \setminus \{0\}$, then*

$$\text{grad}\,R_f(x) = \frac{2}{||x||^2}(f(x) - R_f(x) \cdot x). \tag{11}$$

This is a direct result of (10), since $f^* = f$.

**Corollary 3.2.** *When the linear operator* $f : E \to E$ *is selfadjoint and when* $x \in E \setminus \{0\}$ *is an extreme value point for* $R_f(x)$*, then* $x$ *is an eigenvector for* $f$*.*

*Proof.* If $x$ is an extreme value point on the open set $E \setminus \{0\}$, then it is a critical point for $R_f(x)$ and hence, $\operatorname{grad} R_f(x) = 0$; according to (11), it results that $f(x) = R_f(x) \cdot x$. The corresponding eigenvalue is precisely $R_f(x)$. □

**Corollary 3.3.** *When the linear operator* $f : E \to E$ *is orthogonal and* $x \in E \setminus \{0\}$ *is an extreme local point for* $R_f(x)$*, then the subspace* $G$ *of* $E$ *generated by* $x$ *and* $f(x)$ *is invariant by* $f$*.*

*Proof.* In this case, $f$ is an isomorphism and $f^* = f^{-1}$. According to (10), it follows that $f(x) + f^{-1}(x) - 2R_f(x) \cdot x = 0$. When $z \in G$, then $z = \alpha x + \beta f(x)$, with $\alpha, \beta \in \mathbb{R}$, hence $f(z) = \alpha f(x) + \beta f(f(x))$. But $f(f(x)) + x - 2R_f(x) \cdot f(x) = 0$. Consequently,

$$f(z) = \alpha f(x) + \beta(2R_f(x) \cdot f(x) - x);$$

so $f(z) \in G$. □

## REFERENCES


[1] *M.W. Hirsch and St. Smale*, Differential Equations, Dynamical Systems and Linear Algebra, Academic Press, 1974.

[2] *R.A. Horn and Ch.R. Johnson*, Analiză matricială (Matriceal Analysis), Ed. Theta, 2001.

[3] *A. Niţă*, Generalized inverse of a matrix with applications to optimization of some systems (Ph.D. Thesis, in Romanian), Univ. of Bucharest, Faculty of Mathematics and Computer Sciences, 2004.

[4] *G. Strang*, Linear algebra and its applications, Academic Press, 1976.

[5] *H. Yserentant*, A Short Theory of the Rayleigh-Ritz Method, Computational Methods is Applied Mathematics, **13**, 4, 495-502 (2013).

[6] *L.H. Zhang, J. Xue and R.C. Li*, Rayleigh-Ritz Approximation For The Linear Response Eigenvalue Problem, SIAM J. Matrix Anal. & Appl. **35**, 2, 765-782 (2014).