3.7. **Solving linear systems by minimization.** A scalar equation \(Ax = b\) can be solved by minimization: its solution coincides with the point \(x = x_m\) where the parabola \(p(x) = \frac{1}{2}Ax^2 - bx\) has a minimum (assuming \(A > 0\)). To generalize this minimization solution to higher dimensions, note that the derivative of \(p\) is \(p'(x) = Ax - b\), whose critical point is the solution of the linear equation \(Ax = b\), and the second derivative is \(p''(x) = A > 0\), which ensures that the critical point is a minimum. The idea is then to construct a function \(p(x)\) whose gradient is \(Ax - b\), and Hessian is \(A\):

**Theorem 31.** Let \(A \in M_n(\mathbb{R})\) be a positive definite matrix, and \(b \in \mathbb{R}^n\).

The quadratic form

\[
p(x) = \frac{1}{2} \langle x, Ax \rangle - \langle x, b \rangle
\]

has a global minimum.

The minimum is attained at a point \(x_m\) satisfying \(Ax_m = b\) and the minimum value is \(p(x_m) = -\frac{1}{2} \langle A^{-1}b, b \rangle\).

**Proof:**

The gradient of \(p(x)\) equals \(Ax - b\) since from

\[
p(x) = p(x_1, \ldots, x_n) = \frac{1}{2} \sum_{i,j=1,\ldots,n} A_{ij} x_i x_j - \sum_{i=1,\ldots,n} b_i x_i
\]

we get, using that \(A\) is symmetric,

\[
\frac{\partial p}{\partial x_k} = \frac{1}{2} \sum_{j=1,\ldots,n} A_{kj} x_j + \frac{1}{2} \sum_{i=1,\ldots,n} A_{ik} x_i - b_k = \sum_{j=1,\ldots,n} A_{kj} x_j + b_k = (Ax)_k - b_k
\]

and the Hessian of \(p(x)\) is \(A\) since \(\frac{\partial^2 p}{\partial x_i \partial x_k} = A_{ik}\).

[Here is a more compact way to do this calculation:]

\[
\frac{\partial p}{\partial x_k} = \frac{1}{2} \langle \frac{\partial}{\partial x_k} x, Ax \rangle + \frac{1}{2} \langle x, A \frac{\partial}{\partial x_k} x \rangle - \langle \frac{\partial}{\partial x_k} x, b \rangle
\]

\[
= \frac{1}{2} \langle e_k, Ax \rangle + \frac{1}{2} \langle x, Ae_k \rangle - \langle e_k, b \rangle = \langle e_k, Ax \rangle - b_k = (Ax)_k - b_k
\]

and the Hessian of \(p(x)\) is \(A\) since \(\frac{\partial^2 p}{\partial x_i \partial x_k} = \frac{\partial}{\partial x_k} \langle e_k, Ax \rangle = \langle e_k, A \frac{\partial}{\partial x_j} x \rangle = \langle e_k, Ae_j \rangle = A_{jk}\).

Note that the linear system \(Ax = b\) has a unique solution \(x = x_m = A^{-1}b\) (since \(A\) has no zero eigenvalues).

Write the Taylor expansion of \(p\) at \(x = x_m\), which stops at the second order terms, since \(p\) is a polynomial of degree 2, hence derivatives of higher order vanish:

\[
p(x) = p(x_m) + \langle \nabla p(x_m), (x - x_m) \rangle + \frac{1}{2} \langle (x - x_m), (Hp)(x_m) (x - x_m) \rangle
\]

\[
= p(x_m) + \langle (Ax_m - b), (x - x_m) \rangle + \frac{1}{2} \langle (x - x_m), A (x - x_m) \rangle
\]
where the last inequality holds since \( A \) is positive definite; this means that 
\( p(x_m) \) is an absolute minimum. Furthermore

\[
p(x_m) = \frac{1}{2} \langle x_m, A x_m \rangle - \frac{1}{2} \langle x_m, b \rangle = -\frac{1}{2} \langle A^{-1} b, b \rangle
\]

\( \Box \)

3.8. **Generalized eigenvalue problems.** The following type of problem appears in applications (for example they arise in discretizations of continuous problems):

**Problem:** given two \( n \times n \) matrices \( A \) and \( B \) find the numbers \( \lambda \) so that

\[
Av = \lambda Bv \quad \text{for some } v \neq 0
\]

Clearly, such numbers \( \lambda \) must satisfy the equation \( \det(A - \lambda B) = 0 \), which is a polynomial in \( \lambda \) of degree at most \( n \).

Since \( \det(A - \lambda B) = \lambda^n \det(A - \lambda B) \) wee that the coefficient of \( \lambda^n \) is \( \det(-B) \). Therefore, if \( B \) is not invertible, the degree of \( \det(A - \lambda B) \) is less than \( n \).

On the other hand, if \( B \) is invertible, then (18) is equivalent to \( B^{-1}Av = \lambda v \) so \( \lambda \) and \( v \) are the eigenvalues and eigenvectors of the matrix \( B^{-1}A \).

For real matrices, \( A \) symmetric and \( B \) positive definite, the eigenvalues/vectors of the generalized problem have special properties, which are derived below.

If \( B \) is positive definite and real, then \( B = M^T M \) for a positive definite real matrix \( M \), by Theorem 23 (iv). Equation (18) is \( Av = \lambda M^T Mv \) therefore \( (M^T)^{-1}Av = \lambda Mv \), where denoting \( Mv = y \) and \( C = M^{-1} \) the equation becomes

\[
(C^TAC)y = \lambda y
\]

which is an eigenvalue problem for the symmetric matrix \( C^TAC \): there are \( n \) real eigenvalues, and \( n \) orthonormal eigenvectors \( u_1, \ldots, u_n \). Going back through the substitution, let \( v_j \)'s be so that \( u_j = Mv_j \), we find the eigenvectors of the generalized problem (18) as \( v_1, \ldots, v_n \) where

\[
\delta_{ij} = \langle u_j, u_i \rangle = \langle Mv_j, Mv_i \rangle = \langle v_j, v_i \rangle = \langle v_j, Bv_i \rangle
\]

and therefore

\[
\langle v_j, Bv_i \rangle = \delta_{ij}, \text{ for } i, j = 1, \ldots, n
\]

meaning that the \( v_j \) are \( B \)-orthonormal, i.e. orthonormal with respect to the inner product \( \langle \cdot, \cdot \rangle_B \), see Theorem 23 (iii).

Also

\[
\langle v_j, Av_i \rangle = \langle v_j, \lambda_i Bv_i \rangle = \lambda_i \langle v_j, M^T Mv_i \rangle = \lambda_i \langle v_j, Bv_i \rangle = \lambda_i \delta_{ij}
\]
so
\[ (21) \quad \langle v_j, Av_i \rangle = \delta_{ij} \lambda_i, \quad \text{for } i, j = 1, \ldots, n \]
meaning that the vectors \( v_1, \ldots, v_n \) are also \( A \)-orthogonal.

In matrix notation, if \( S = [v_1, \ldots, v_n] \) is the matrix with columns the
generalized eigenvectors, then (21) is \( S^T AS = \Lambda \), and (20) is \( S^T BS = I \).

The matrices \( A \) and \( B \) are simultaneously diagonalized by a congruence
transformation! We proved:

**Theorem 32.** Consider the real matrices: \( A \) symmetric and \( B \) positive
definite. Then the quadratic forms

\[ \begin{align*}
q(x) &= \langle x, Ax \rangle, \quad r(x) = \langle x, Bx \rangle
\end{align*} \]

are simultaneously diagonalizable.

More precisely, there exists an invertible matrix \( S \) so that \( S^T AS = \Lambda \) and
\( S^T BS = I \) and if \( x = Sy \) then

\[ \begin{align*}
q(x) &= q(Sy) = \langle y, \Lambda y \rangle, \quad r(x) = r(Sy) = \langle y, y \rangle
\end{align*} \]

The diagonal matrix \( \Lambda \) consists of the generalized eigenvalues solutions of
the problem \( Av = \lambda Bv \) and the columns of \( S \) are its generalized eigenvectors.
4. The Rayleigh’s principle. The minimax theorem for the eigenvalues of a self-adjoint matrix

Eigenvalues of self-adjoint matrices are easy to calculate. This section shows how this is done using a minimization, or maximization procedure.

4.1. The Rayleigh’s quotient.

Definition 33. Let $A = A^*$ be a self-adjoint matrix. The Rayleigh’s quotient is the function

$$R(x) = \frac{\langle x, Ax \rangle}{\|x\|^2}, \text{ for } x \neq 0$$

Note that

$$R(x) = \frac{x}{\|x\|}, A \frac{x}{\|x\|} = \langle u, Au \rangle \text{ where } u = \frac{x}{\|x\|}$$

so in fact, it suffices to define the Rayleigh’s quotient on unit vectors.

The set of unit vectors in $\mathbb{R}^n$ (or in $\mathbb{C}^n$), is called the $n - 1$ dimensional sphere in $\mathbb{R}^n$ (or in $\mathbb{C}^n$):

$$S^{n-1}_F = \left\{ u \in F^n \mid \|u\| = 1 \right\}$$

For example, the sphere in $\mathbb{R}^2$ is the unit circle (it is a curve, it has dimension 1), the sphere in $\mathbb{R}^3$ is the unit sphere (it is a surface, it has dimension 2); for higher dimensions we need to use our imagination.

4.2. Extrema of the Rayleigh’s quotient.

4.2.1. Closed sets, bounded sets, compact sets. You probably know very well the extreme value theorem for continuous function on the real line:

Theorem 34. The extreme value theorem in dimension one.

A function $f(x)$ which is continuous on a closed and bounded interval $[a, b]$ has a maximum value (and a minimum value) on $[a, b]$.

To formulate an analogue of this theorem in higher dimensions we need to specify what we mean by a closed set and by a bounded set.

Definition 35. A set $S$ is called closed if it contains all its limit points: if a sequence of points in $S$, $\{x_k\}_k \subset S$ converges, $\lim_{k \to \infty} x_k = x$, then the limit $x$ is also in $S$.

For example, the intervals $[2, 6]$ and $[2, +\infty)$ are closed in $\mathbb{R}$, but $[2, 6)$ is not closed. The closed unit disk $\{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$ is closed in $\mathbb{R}^2$, but the punctured disk $\{x \in \mathbb{R}^2 \mid 0 < \|x\| \leq 1\}$ or the open disk $\{x \in \mathbb{R}^2 \mid \|x\| < 1\}$ are not closed sets.

Definition 36. A set $S$ is called bounded if there is a number larger than all the lengths of the vectors in $S$: there is $M > 0$ so that $\|x\| \leq M$ for all $x \in S$. 
For example, the intervals \([2, 6]\) and \([2, 6)\) are bounded in \(\mathbb{R}\), but \([2, +\infty)\) is not. The square \(\{x \in \mathbb{R}^2 \mid |x_1| < 1, \text{ and } |x_2| < 1\}\) is bounded in \(\mathbb{R}^2\), but the vertical strip \(\{x \in \mathbb{R}^2 \mid |x_1| < 1\}\) is not.

**Theorem 37. The extreme value theorem in finite dimensions.**

A function \(f(x)\) which is continuous on a closed and bounded set \(S\) in \(\mathbb{R}^n\) or \(\mathbb{C}^n\) has a maximum value (and a minimum value) on \(S\).

In infinite dimensions Theorem 37 is not true in this form. A more stringent condition on the set \(S\) is needed to ensure existence of extreme values of continuous functions on \(S\) (the set must be compact).

It is intuitively clear (and rigorously proved in mathematical analysis) that any sphere in \(F^n\) is a closed and bounded set.

### 4.2.2. Minimum and maximum of the Rayleigh’s quotient.

The Rayleigh’s quotient calculated on unit vectors is a quadratic polynomial, and therefore

**Proposition 38. The Rayleigh’s quotient has a maximum and a minimum.**

What happens if \(A\) is not self-adjoint? Recall that the quadratic form \(\langle x, Ax \rangle\) has the same value if we replace \(A\) by its self-adjoint part, \(\frac{1}{2} (A + A^*)\), therefore the Rayleigh’s quotient of \(A\) is the same as the Rayleigh’s quotient of the self-adjoint part of \(A\) (information about \(A\) is lost).

The extreme values of the Rayleigh’s quotient are linked to the eigenvalues of the self-adjoint matrix \(A\). To see this, diagonalize the quadratic form \(\langle x, Ax \rangle\): consider a unitary matrix \(U\) which diagonalizes \(A\):

\[
U^*AU = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)
\]

In the new coordinates \(y = U^*x\) we have

\[
\langle x, Ax \rangle = \langle x, UAU^*x \rangle = \langle U^*x, \Lambda U^*x \rangle = \langle y, \Lambda y \rangle = \sum_{j=1}^{n} \lambda_j |y_j|^2
\]

which together with \(\|x\| = \|Uy\| = \|y\|\) give

\[
(22) \quad R(x) = R(Uy) = \sum_{j=1}^{n} \lambda_j |y_j|^2 = \sum_{j=1}^{n} \lambda_j \frac{|y_j|^2}{\|y\|^2} = R_U(y)
\]

Since \(A\) is self-adjoint, its eigenvalues \(\lambda_j\) are real; assume them ordered in an increasing sequence:

\[
\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n
\]

Then clearly

\[
\sum_{j=1}^{n} \lambda_j |y_j|^2 \leq \sum_{j=1}^{n} \lambda_n |y_j|^2 = \lambda_n \|y\|^2
\]
and
\[ \sum_{j=1}^{n} \lambda_j |y_j|^2 \geq \lambda_1 \sum_{j=1}^{n} |y_j|^2 = \lambda_1 \|y\|^2 \]
therefore
\[ \lambda_1 \leq R(x) \leq \lambda_n \quad \text{for all } x \neq 0 \]
Equalities are attained since \( R_U(e_1) = 1 \) and \( R_U(e_n) = n \). Going to coordinates \( x \) minimum is attained for \( x = Ue_1 = v_1 \) = eigenvector corresponding to \( \lambda_1 \) since \( R(v_1) = R_U(e_1) = \lambda_1 \), and for \( x = Uv_n = v_n \) = eigenvector corresponding to \( \lambda_n \), maximum is attained since \( R(v_n) = R_U(e_n) = \lambda_n \). This proves:

**Theorem 39.** If \( A \) is a self-adjoint matrix then
\[
\max \frac{\langle x, Ax \rangle}{\|x\|^2} = \lambda_n \text{ the max eigenvalue of } A, \text{ attained for } x = v_n
\]
and
\[
\min \frac{\langle x, Ax \rangle}{\|x\|^2} = \lambda_1 \text{ the min eigenvalue of } A, \text{ attained for } x = v_1
\]
As an important consequence in numerical calculations: the maximum eigenvalue of \( A \) can be found by solving a maximization problem, and the minimum eigenvalue - by a minimization problem.

4.3. The minimax principle. Reducing the dimension of \( A \) we can find all the eigenvalues, one by one. This reduction of the dimension is done using:

4.3.1. Splitting of the space under the action of a self-adjoint matrix. Any matrix leaves invariant its eigenspaces. Since eigenvectors of a self-adjoint matrix \( A \) form an orthogonal set, then self-adjoint matrices leaves invariant their orthogonal complement as well:

**Remark.** Let \( A \) be any \( n \times n \) self-adjoint matrix.

a) The vector space \( F^n \) splits
\[ F^n = \oplus_{\lambda \in \sigma(A)} V_{\lambda} \]
where each eigenspace \( V_{\lambda} \) is invariant under \( A \), and so is its orthogonal complement, which equals the direct sum of all the other eigenspaces:
\[ V_{\lambda}^\perp = \oplus_{\lambda' \in \sigma(A), \lambda' \neq \lambda} V_{\lambda'} \]
b) If \( v \) is an eigenvector then \( A \) leaves invariant \( Sp(v) \) and \( Sp(v)^\perp \).

**Proof:**
Only part b) needs a proof.
\[ A(Sp(v)) \subset Sp(v) \text{ because if } x \in Sp(v) \text{ then } x = xv \text{ therefore } Ax = cAv = c\lambda v \in Sp(v). \]
\[ A(Sp(v)^\perp) \subset Sp(v)^\perp \text{ because if } y \in Sp(v)^\perp \text{ then this means that } \langle y, v \rangle = 0. \]
Then \( \langle Ay, v \rangle = \langle y, A^*v \rangle = \langle y, Av \rangle = \langle y, \lambda v \rangle = \lambda \langle y, v \rangle = 0. \) \( \Box \)
4.3.2. Minimax and maximin. We saw that \( \max R(x) = \lambda_n = R(v_n) \). Then the matrix \( A \), as a linear transformation of the \( n-1 \) dimensional vector space \( Sp(v_n) \) to itself has its largest eigenvalue \( \lambda_{n-1} \) (we reduced the dimension!). Then

\[
\max_{x \in Sp(v_n)} R(x) = \lambda_{n-1}
\]
is attained for \( x = v_{n-1} \).

Note that the statement \( x \in Sp(v_n) \) can be formulated as the constraint \( \langle x, v_n \rangle = 0 \):

\[
\max_{\langle x, v_n \rangle = 0} R(x) = \lambda_{n-1}
\]

We can do even better: we can obtain \( \lambda_{n-1} \) without knowing \( v_n \) or \( \lambda_n \).

To achieve this, subject \( x \) to any constraint:

\[
\langle x, z \rangle = 0 \text{ for some } z \neq 0.
\]

It is easier to see what happens in coordinates \( y = U^*x \) in which \( A \) is diagonal. The constraint \( \langle x, z \rangle = 0 \) is equivalent to \( \langle y, w \rangle = 0 \) where \( w = Uz \) is some nonzero vector. On one hand, obviously

\[
\max_{y: \langle y, w \rangle = 0} R_U(y) \leq \lambda_n
\]

(23)

\[
\max_{y: \langle y, w \rangle = 0} R_U(y) \geq \lambda_{n-1}
\]

which implies that

(24)

\[
\min_{w \neq 0} \max_{y: \langle y, w \rangle = 0} R_U(y) \geq \lambda_{n-1}
\]

We now argue that equality is attained in (24) for special \( w \) and \( y \). Indeed, consider a nonzero vector \( \tilde{y} = (0, \ldots, 0, y_{n-1}, y_n)^T \) with \( \langle \tilde{y}, w \rangle = 0 \). \( \langle \tilde{y}, w \rangle \) is easy to find: if \( w_n \neq 0 \) take \( y_{n-1} = 1 \) and \( y_n = -w_{n-1}/w_n \), and if \( w_n = 0 \) take \( y_{n-1} = 0, y_n = 1 \).

Using formula (22)

\[
R_U(\tilde{y}) = \frac{\lambda_{n-1}|y_{n-1}|^2 + \lambda_n|y_n|^2}{|y_{n-1}|^2 + |y_n|^2} \geq \frac{\lambda_{n-1}|y_{n-1}|^2 + \lambda_{n-1}|y_n|^2}{|y_{n-1}|^2 + |y_n|^2} = \lambda_{n-1}
\]

Since for \( w = e_n \) we have equality:

\[
\max_{y: \langle y, e_n \rangle = 0} R_U(y) = \lambda_{n-1}
\]

then in (24) there is equality

\[
\min_{w \neq 0} \max_{y: \langle y, w \rangle = 0} R_U(y) = \lambda_{n-1}
\]

In a similar way it is shown that \( \lambda_{n-2} \) is obtained by a minimum-maximum process, but with two constraints:

(25)

\[
\min_{w_1, w_2 \neq 0} \max_{\langle y, w_1 \rangle = 0, \langle y, w_2 \rangle = 0} R_U(y) = \lambda_{n-2}
\]

\[
R_U(y) = \lambda_{n-2}
\]
Indeed, consider a nonzero vector \( \tilde{y} = (0, \ldots, 0, y_{n-2}, y_{n-1}, y_n)^T \) satisfying \( \langle \tilde{y}, w_1 \rangle = 0 \) and \( \langle \tilde{y}, w_2 \rangle = 0 \). Then in formula (22)

\[
R_U(\tilde{y}) = \frac{\lambda_{n-2}|y_{n-2}|^2 + \lambda_{n-1}|y_{n-1}|^2 + \lambda_n|y_n|^2}{|y_{n-2}|^2 + |y_{n-1}|^2 + |y_n|^2} \\
\geq \frac{\lambda_{n-2}|y_{n-2}|^2 + \lambda_{n-2}|y_{n-1}|^2 + \lambda_{n-2}|y_n|^2}{|y_{n-2}|^2 + |y_{n-1}|^2 + |y_n|^2} = \lambda_{n-2}
\]

which shows that

\[
\max_{\langle y, w_1 \rangle = 0 \atop \langle y, w_2 \rangle = 0} R_U(y) \geq \lambda_{n-2}
\]

Since for \( w_1 = e_n \) and \( w_2 = e_{n-1} \) we have equality in (26), this implies (25).

Step by step, adding one extra constraint, the minimax procedure yields the next largest eigenvalue.

Going back to the variable \( x \) it is found that:

**Theorem 40. The minimax principle**

*Let \( A \) be a self-adjoint matrix, with its eigenvalues numbered in an increasing sequence:

\[ \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \]

*corresponding to the eigenvectors \( v_1, \ldots, v_n \).

Then its Rayleigh's quotient

\[
R(x) = \frac{\langle x, Ax \rangle}{\|x\|^2}
\]

satisfies

\[
\max_{x \neq 0} R(x) = \lambda_n \\
\min_{z \neq 0} \max_{\langle x, z \rangle = 0} R(x) = \lambda_{n-1} \\
\min_{z_1, z_2 \neq 0} \max_{\langle x, z_1 \rangle = 0, \langle x, z_2 \rangle = 0} R(x) = \lambda_{n-2} \\
\text{and in general}
\]

\[
\min_{z_1, \ldots, z_k \neq 0} \max_{\langle x, z_1 \rangle = 0, \ldots, \langle x, z_k \rangle = 0} R(x) = \lambda_{n-k}, \quad k = 1, 2, \ldots, n - 1
\]

**Remark.** Sometimes the minimax principle is formulated as

\[
\min_{V_j} \max_{x \in V_j} R(x) = \lambda_j, \quad j = 1, \ldots, n
\]

where \( V_j \) denotes an arbitrary subspace of dimension \( j \).
The two formulations are equivalent since the set
\[ V_{n-k} = \{ \mathbf{x} \mid \langle \mathbf{x}, \mathbf{z}_1 \rangle = 0, \ldots, \langle \mathbf{x}, \mathbf{z}_k \rangle = 0 \} \]
is a vector space of dimension \( n - k \) if \( \mathbf{z}_1, \ldots, \mathbf{z}_k \) are linearly independent.

A similar construction starting with the lowest eigenvalue produces:

**Theorem 41. The maximin principle**

Under the assumptions of Theorem 40

\[ \min_{\mathbf{x} \neq 0} R(\mathbf{x}) = \lambda_1 \]
\[ \max_{\mathbf{z} \neq 0} \min_{\langle \mathbf{x}, \mathbf{z} \rangle = 0} R(\mathbf{x}) = \lambda_2 \]

and in general

\[ \max_{\mathbf{z}_1, \ldots, \mathbf{z}_k \neq 0} \min_{\langle \mathbf{x}, \mathbf{z}_1 \rangle = 0} \ldots \min_{\langle \mathbf{x}, \mathbf{z}_k \rangle = 0} R(\mathbf{x}) = \lambda_{k+1}, \quad k = 1, 2, \ldots, n - 1 \]

4.4. The minimax principle for the generalized eigenvalue problem.

Suppose \( \lambda_1 \leq \lambda_1 \leq \ldots \leq \lambda_n \) are eigenvalues for the problem

\[ (27) \quad A\mathbf{v} = \lambda B\mathbf{v}, \quad A \text{ symmetric}, \quad B \text{ positive definite} \]

It was seen in §3.8 that if \( S = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \) is the matrix whose columns are the generalized eigenvectors of the problem (27), then both matrices \( A \) and \( B \) are diagonalized using a congruence transformation: \( S^T A S = \Lambda \) and \( S^T B S = I \).

Defining

\[ R(\mathbf{x}) = \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, B\mathbf{x} \rangle} \]

it is found that in coordinates \( \mathbf{x} = S\mathbf{y} \):

\[ R(\mathbf{x}) = R(S\mathbf{y}) = \frac{\langle S\mathbf{y}, A S\mathbf{y} \rangle}{\langle S\mathbf{y}, B S\mathbf{y} \rangle} = \frac{\langle \mathbf{y}, S^T A S \mathbf{y} \rangle}{\langle \mathbf{y}, S^T B S \mathbf{y} \rangle} = \frac{\lambda_1 y_1^2 + \ldots + \lambda_n y_n^2}{y_1^2 + \ldots + y_n^2} \]

and therefore

\[ \max R(\mathbf{x}) = \lambda_n, \quad \min R(\mathbf{x}) = \lambda_1 \]