A constrained eigenvalue problem

Report

Author(s):
Gander, Walter; Golub, Gene Howard; Matt, Urs von

Publication date:
1988

Permanent link:
https://doi.org/10.3929/ethz-a-000486121

Originally published in:
ETH, Eidgenössische Technische Hochschule Zürich, Institut für Informatik, Fachgruppe Wissenschaftliches Rechnen 92

Eidgenössische
Technische Hochschule
Zürich

Institut für Informatik
Fachgruppe
Wissenschaftliches Rechnen

Walter Gander
Gene Golub
Urs von Matt

A Constrained Eigenvalue Problem

October 1988

92

88. 11. 5

Authors' addresses:

Institut für Informatik
ETH-Zentrum
CH-8092 Zürich / Switzerland

Department of Computer Science
Stanford University
Stanford, California 94305 / USA

1 INTRODUCTION In this paper we consider the following mathematical and computational problem. Given the quantities

A: $(n+m)$-by-$(n+m)$ matrix, symmetric, $n > 0$
N: $(n+m)$-by-m matrix with full rank
$\mathbf{t}$: vector of dimension m with $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\| < 1$

Determine an $\mathbf{x}$ such that

$$\mathbf{x}^{\mathrm{T}}\mathrm{A}\mathbf{x} = \min \tag{1}$$

subject to the constraints

$$\mathrm{N}^{\mathrm{T}}\mathbf{x} = \mathbf{t} \tag{2}$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{x} = 1. \tag{3}$$

Variants of this problem occur in many applications [1,5,7,8,11]. The problem has been studied previously when $\mathbf{t} = \mathbf{0}$, the null vector, (cf. [4,6]).

When $\mathbf{t} \neq \mathbf{0}$, then the problem becomes more complicated. We now motivate our assumptions. Suppose N does not have full rank. If $\mathbf{t}$ is not in the range of N^{T} the problem has no solution. If, however, the linear constraints are consistent, we can deflate the system until we get sub-matrix of full rank. In the extreme case, where $\mathrm{N} = 0$ and $\mathbf{t} = \mathbf{0}$, the problem reduces to the ordinary eigenvalue problem with an eigenvector corresponding to the smallest eigenvalue of A as the solution.

Now consider the quantity $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\|$. As $(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}$ denotes the unique solution of $\mathrm{N}^{\mathrm{T}}\mathbf{x} = \mathbf{t}$ of minimal norm, we can make the following distinctions: when $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\| > 1$ there is no solution. In the case of $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\| = 1$ the unique solution is given by $\mathbf{x} = (\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}$. Therefore, the condition $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\| < 1$ is the only interesting case.

Thus when $\|(\mathrm{N}^{\mathrm{T}})^{+}\mathbf{t}\| < 1$, we can always find an $\mathbf{x}$ that satisfies both constraints and hence there always exists at least one solution to the problem.

2 PROBLEM SIMPLIFICATION

In this section we will normalize the problem in order to study its solvability. These considerations are not only of theoretical interest, but they also serve as a basis for the numerical calculations.

2.1 Elimination of the Linear Constraint

In the domain of the transformation N^T we can distinguish two fundamental subspaces:

$$\mathcal{N}(\mathrm{N}^\mathrm{T}) \;\text{: nullity of } \mathrm{N}^\mathrm{T}$$

$$\mathcal{N}(\mathrm{N}^\mathrm{T})^\perp\text{: orthogonal complement.}$$

If the transformation N^T is being restricted on $\mathcal{N}(\mathrm{N}^\mathrm{T})^\perp$ it acts as a bijection between $\mathcal{N}(\mathrm{N}^\mathrm{T})^\perp$ and $\mathcal{R}(\mathrm{N}^\mathrm{T})$ with the inverse $(\mathrm{N}^\mathrm{T})^+$. Thus the general solution of the first constraint is given by

$$\mathbf{x} = (\mathrm{N}^\mathrm{T})^+\mathbf{t} + \xi$$

with an arbitrary $\xi \in \mathcal{N}(\mathrm{N}^\mathrm{T})$. Now using the singular value decomposition of N^T,

$$\mathrm{N}^\mathrm{T} = \mathrm{U}\Sigma\mathrm{V}^\mathrm{T}$$

with

$$\mathrm{U} = \begin{array}{cc} m & \\ \left[\begin{array}{c} \\ \quad\mathrm{U}\quad \\ \\ \end{array}\right] & m \end{array}$$

$$\Sigma = \begin{array}{cc} \begin{array}{cc} \qquad m \qquad & n \end{array} & \\ \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_m & \end{array}\right] & m \end{array}$$

$$\mathrm{V} = \begin{array}{cc} \begin{array}{cc} \qquad m \qquad & n \end{array} & \\ \left[\begin{array}{c|c} \\ \qquad\mathrm{V}_1\qquad & \mathrm{V}_2 \\ \\ \end{array}\right] & m+n \quad , \end{array}$$

we can write $\mathbf{x}$ as

$$\mathbf{x} = \mathrm{V}\Sigma^+\mathrm{U}^\mathrm{T}\mathbf{t} + \mathrm{V}_2\mathbf{z}$$

with an arbitrary $\mathbf{z}$. For the subsequent minimization, we only have to consider the second constraint (3).

For practical reasons, we can use the QR-decomposition of N instead of the singular value decomposition. To simplify the first constraint (2), we write

$$(4)\qquad \mathrm{P^T N} = \begin{bmatrix} \mathrm{R} \\ 0 \end{bmatrix}$$

where P denotes an orthogonal matrix, and R is a m-by-m upper triangular matrix. Now the problem can be formulated as

$$\mathbf{x}^{\mathrm{T}}\mathrm{A}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathrm{PP^TAPP^T}\mathbf{x} = \min$$

$$\mathrm{N^T}\mathbf{x} = \begin{bmatrix} \mathrm{R^T} & 0 \end{bmatrix} \mathrm{P^T}\mathbf{x} = \mathbf{t}$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathrm{PP^T}\mathbf{x} = 1.$$

We now make the definitions,

$$(5)\qquad \mathrm{P^TAP} =: \begin{array}{cc} \begin{array}{cc} m & n \end{array} & \\ \begin{bmatrix} \mathrm{B} & \Gamma^{\mathrm{T}} \\ \Gamma & \mathrm{C} \end{bmatrix} & \begin{array}{c} m \\ n \end{array} \end{array} \quad ;$$

$$(6)\qquad \mathrm{P^T}\mathbf{x} =: \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \begin{array}{c} m \\ n \end{array} \quad .$$

Note that $\mathrm{C^T} = \mathrm{C}$. Now

$$\begin{aligned} \mathbf{x}^{\mathrm{T}}\mathrm{A}\mathbf{x} &= \begin{bmatrix} \mathbf{y}^{\mathrm{T}} & \mathbf{z}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathrm{B} & \Gamma^{\mathrm{T}} \\ \Gamma & \mathrm{C} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} \\ &= \mathbf{y}^{\mathrm{T}}\mathrm{B}\mathbf{y} + \mathbf{y}^{\mathrm{T}}\Gamma^{\mathrm{T}}\mathbf{z} + \mathbf{z}^{\mathrm{T}}\Gamma\mathbf{y} + \mathbf{z}^{\mathrm{T}}\mathrm{C}\mathbf{z} \\ &= \mathbf{y}^{\mathrm{T}}\mathrm{B}\mathbf{y} + 2\mathbf{z}^{\mathrm{T}}\Gamma\mathbf{y} + \mathbf{z}^{\mathrm{T}}\mathrm{C}\mathbf{z}, \end{aligned}$$

$$\mathrm{N^T}\mathbf{x} = \begin{bmatrix} \mathrm{R^T} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} = \mathrm{R^T}\mathbf{y} = \mathbf{t},$$

$$\mathbf{x}^{\mathrm{T}}\mathbf{x} = \begin{bmatrix} \mathbf{y}^{\mathrm{T}} & \mathbf{z}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} = \mathbf{y}^{\mathrm{T}}\mathbf{y} + \mathbf{z}^{\mathrm{T}}\mathbf{z} = 1.$$

So we have reduced the constraint (2) to an ordinary linear system with an upper triangular matrix. Now

$$\mathbf{y} = \mathrm{R}^{-\mathrm{T}}\mathbf{t} \tag{7}$$

and with the help of the definitions

$$s^2 := 1 - \mathbf{y}^{\mathrm{T}}\mathbf{y} > 0 \tag{8}$$

$$\mathbf{b} := -\Gamma\mathbf{y}, \tag{9}$$

we get the simplified problem

$$\begin{aligned} \mathbf{z}^{\mathrm{T}}\mathrm{C}\mathbf{z} - 2\mathbf{b}^{\mathrm{T}}\mathbf{z} &= \min \\ \mathbf{z}^{\mathrm{T}}\mathbf{z} &= s^2. \end{aligned} \tag{10}$$

This problem has been extensively studied in the literature (cf. [2, 7,8,10]).

2.2 Stationary Points In order to calculate the stationary points, we set up the so-called *Lagrange principal function*:

$$\Phi(\mathbf{z}, \lambda) := \mathbf{z}^{\mathrm{T}}\mathrm{C}\mathbf{z} - 2\mathbf{b}^{\mathrm{T}}\mathbf{z} - \lambda(\mathbf{z}^{\mathrm{T}}\mathbf{z} - s^2) \tag{11}$$

Differentiating Φ by $\mathbf{z}$ and λ yields the equations

$$\begin{aligned} 2\mathrm{C}\mathbf{z} - 2\mathbf{b} - 2\lambda\mathbf{z} &= 0 \\ \mathbf{z}^{\mathrm{T}}\mathbf{z} - s^2 &= 0 \end{aligned}$$

or normalized

$$\begin{aligned} \mathrm{C}\mathbf{z} &= \lambda\mathbf{z} + \mathbf{b} \\ \mathbf{z}^{\mathrm{T}}\mathbf{z} &= s^2. \end{aligned} \tag{12}$$

Now let us compare the values $\mathbf{z}^{\mathrm{T}}\mathrm{C}\mathbf{z} - 2\mathbf{b}^{\mathrm{T}}\mathbf{z}$ of different tuples $(\lambda, \mathbf{z})$. Following the proof given in [3,12], a short calculation shows that the smallest λ is needed in order to minimize the value

$\mathbf{z}^T C\mathbf{z} - 2\mathbf{b}^T\mathbf{z}$. So in place of the original minimization we can solve the *Lagrange equations*

$$
\begin{aligned}
C\mathbf{z} &= \lambda\mathbf{z} + \mathbf{b} \\
\mathbf{z}^T\mathbf{z} &= s^2 \\
\lambda &= \min .
\end{aligned}
\tag{13}
$$

3 SOLVABILITY We will now investigate the solvability of the Lagrange equations (13). Simultaneously this analysis will point out a first method to solve the problem.

3.1 Explicit Secular Equation For our discussion, we need the eigenvalue decomposition

$$
C = QDQ^T \tag{14}
$$

where

$$
D = \operatorname{diag}(\delta_1, \dots, \delta_n) \qquad \delta_1 \le \delta_2 \le \cdots \le \delta_n
$$

and

$$
Q^T Q = I.
$$

Thus the Lagrange equations (13) are transformed as follows:

$$
\begin{aligned}
QDQ^T\mathbf{z} &= \lambda QQ^T\mathbf{z} + \mathbf{b} \\
\mathbf{z}^T\mathbf{z} &= \mathbf{z}^T QQ^T\mathbf{z} = s^2 .
\end{aligned}
$$

With the definitions

$$
\mathbf{u} := Q^T\mathbf{z} \tag{15}
$$

$$
\mathbf{d} := Q^T\mathbf{b}, \tag{16}
$$

this can be simplified to

$$
\begin{aligned}
D\mathbf{u} &= \lambda\mathbf{u} + \mathbf{d} \\
\mathbf{u}^T\mathbf{u} &= s^2 \\
\lambda &= \min .
\end{aligned}
\tag{17}
$$

First let us suppose $\lambda \in \lambda(\mathrm{D})$. Then there exist diagonal elements δ_i with $\delta_i = \lambda$. For the ensuing discussion the following index sets turn out to be useful:

$$\begin{aligned} \mathrm{I} &:= \{ i \mid \delta_i = \lambda \} \\ \bar{\mathrm{I}} &:= \{ i \mid \delta_i \neq \lambda \} = \{1, \ldots, n\} \setminus \mathrm{I}. \end{aligned}$$

If there exists a corresponding $\mathbf{u}$ for such a λ, it must hold that $\forall i : \ \delta_i u_i = \lambda u_i + d_i$ with $\mathbf{u}^{\mathrm{T}}\mathbf{u} = s^2$. Then for $i \in \mathrm{I}$, it must be true that $\lambda u_i = \lambda u_i + d_i$ and this implies $d_i = 0$. And for $i \in \bar{\mathrm{I}}$, it must hold that:

$$u_i = \frac{d_i}{\delta_i - \lambda}.$$

The normalization condition $\mathbf{u}^{\mathrm{T}}\mathbf{u} = s^2$ can only be satisfied if

$$\sum_{i \in \bar{\mathrm{I}}} u_i^2 = \sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 \leq s^2.$$

As a result we have the following three possibilities:

1. There exists no solution $\mathbf{u}$ for a given $\lambda \in \lambda(\mathrm{D})$ if $\exists i \in \mathrm{I} : d_i \neq 0$, or

$$\sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 > s^2.$$

2. There exists a unique solution $\mathbf{u}$ for a given $\lambda \in \lambda(\mathrm{D})$ if $\forall i \in \mathrm{I} : \ d_i = 0$ and

$$\sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 = s^2.$$

   Then $\mathbf{u}$ can be calculated as

$$u_i = \begin{cases} \frac{d_i}{\delta_i - \lambda} & i \in \bar{\mathrm{I}} \\ 0 & i \in \mathrm{I} \end{cases}.$$

3. There exist several solutions **u** for a given $\lambda \in \lambda(\mathrm{D})$ if $\forall i \in \mathrm{I}: \ d_i = 0$, and
$$\sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 < s^2.$$
The possible values for **u** are given by
$$u_i = \frac{d_i}{\delta_i - \lambda} \qquad i \in \bar{\mathrm{I}}$$
$$\sum_{i \in \mathrm{I}} u_i^2 = s^2 - \sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 > 0.$$
Thus the set of all solutions **u** constitutes a manifold of dimension $|\mathrm{I}| - 1$ because the u_i with $i \in \mathrm{I}$ can be chosen arbitrarily on the given hypersphere.

The second case is given by $\lambda \notin \lambda(\mathrm{D})$. Then the inverse $(\mathrm{D} - \lambda \mathrm{I})^{-1}$ exists and **u** has the representation

$$\mathbf{u} = (\mathrm{D} - \lambda \mathrm{I})^{-1} \mathbf{d}. \tag{18}$$

For **u** to solve the normalization condition $\mathbf{u}^{\mathrm{T}} \mathbf{u} = s^2$ it must hold that
$$\mathbf{u}^{\mathrm{T}} \mathbf{u} = \mathbf{d}^{\mathrm{T}} (\mathrm{D} - \lambda \mathrm{I})^{-2} \mathbf{d} = \sum_{i=1}^{n} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 = s^2.$$
We define
$$f(\lambda) := \sum_{i=1}^{n} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 - s^2 \tag{19}$$
as the so-called *explicit secular function* (see Figure 1). Thereby the Lagrange equations (13) have a unique solution **u** for a given $\lambda \notin \lambda(\mathrm{D})$ if and only if the *explicit secular equation*

$$f(\lambda) := \sum_{i=1}^{n} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 - s^2 = 0 \tag{20}$$

is satisfied.

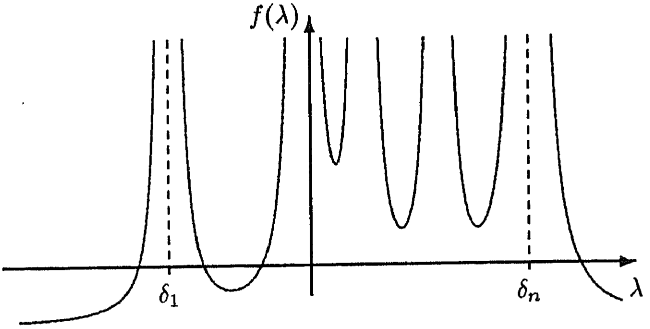

Figure 1: Graph of the secular function

If $d_i = 0$ for all i, the secular function (20) degenerates into $f(\lambda) \equiv -s^2 < 0$ and therefore possesses no solutions. In this case the desired λ lies in the spectrum $\lambda(\mathrm{D})$. Thus let k be the index of the first $d_i \neq 0$, i.e. $d_k \neq 0$ and $\forall i < k : \ d_i = 0$. So we can write the secular function (19) as

$$f(\lambda) = \sum_{i=k}^{n} \left(\frac{d_i}{\delta_i - \lambda} \right)^2 - s^2$$

with $\delta_k \neq 0$. For λ increasing from $-\infty$ to δ_k, $f(\lambda)$ increases strictly since the derivative

$$f'(\lambda) = \sum_{i=k}^{n} \frac{2d_i^2}{(\delta_i - \lambda)^3}$$

is positive for $-\infty < \lambda < \delta_k$. From the limits

$$\lim_{\lambda \to -\infty} f(\lambda) = -s^2$$

$$\lim_{\lambda \to \delta_k -} f(\lambda) = +\infty,$$

it immediately follows that for $\lambda < \delta_k$ there exists exactly one solution.

Then the desired smallest λ can be located either inside or outside the spectrum $\lambda(\mathrm{D})$. With the help of the index set $\mathrm{I} := \{\, i \mid \delta_i = \delta_1 \}$ we can distinguish two alternatives:

1. Either it holds that $\forall i \in \mathrm{I}:\ d_i = 0$, and

   $$\sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \delta_1} \right)^2 \leq s^2.$$

   Whence it follows that for $\lambda = \delta_1$, there exists a solution of the Lagrange equations (13). $f(\lambda)$ possesses no more solutions for $\lambda < \delta_1$ since

   $$f(\delta_1) = \sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \delta_1} \right)^2 - s^2 \leq 0.$$

   Thus with $\lambda = \delta_1$ we have found the smallest λ.

2. Or it holds that $\exists i \in \mathrm{I}:\ d_i \neq 0$, or

   $$\sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \delta_1} \right)^2 > s^2.$$

   It follows that $f(\lambda)$ has a singularity for $\lambda = \delta_1$, or else

   $$f(\delta_1) = \sum_{i \in \bar{\mathrm{I}}} \left(\frac{d_i}{\delta_i - \delta_1} \right)^2 - s^2 > 0.$$

   Therefore $f(\lambda)$ has exactly one solution for $\lambda < \delta_1$. This solution represents the desired smallest λ.

Hence, in both alternatives, the smallest λ always satisfies the condition $\lambda \leq \delta_1$.

3.2 Implicit Secular Equation The above discussion on the location of the smallest λ, that solves the Lagrange equations (13), can be carried out even without the calculation of the eigenvalue decomposition (14) of C. This is useful, when we want to avoid this factorization numerically.

As indicated before, we know that the desired λ satisfies $\lambda \leq \delta_1$. So for $\lambda = \delta_1$ the following cases can be distinguished:

1. The equation $\mathrm{C}\mathbf{z} = \delta_1 \mathbf{z} + \mathbf{b}$ can be inconsistent, i.e.,

   $$(\mathrm{C} - \delta_1 \mathrm{I})(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b} \neq \mathbf{b}.$$

   In this case we have $\lambda < \delta_1$, and the *implicit secular equation*

   $$f(\lambda) = \mathbf{b}^{\mathrm{T}} (\mathrm{C} - \lambda \mathrm{I})^{-2} \mathbf{b} - s^2 = 0 \tag{21}$$

   must be solved.

2. Now we assume that the equation $\mathrm{C}\mathbf{z} = \delta_1 \mathbf{z} + \mathbf{b}$ is consistent. The expression $(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b}$ represents the solution with smallest norm. If $\|(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b}\| > s$, the normalization condition cannot be satisfied, and we have to solve again the secular equation (21).

3. If however $\|(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b}\| = s$, we have found the unique solution of the Lagrange equations (13).

4. Finally it can happen that $\|(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b}\| < s$. Let $\xi^{(1)}, \ldots, \xi^{(k)}$ denote the k orthonormal eigenvectors corresponding to the eigenvalue δ_1. Since

   $$(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b} \perp \mathcal{N}(\mathrm{C} - \delta_1 \mathrm{I}),$$
   $$\xi^{(i)} \in \mathcal{N}(\mathrm{C} - \delta_1 \mathrm{I}),$$

   every vector

   $$\mathbf{z} = (\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b} + c_1 \xi^{(1)} + \cdots + c_k \xi^{(k)}$$

   with

   $$c_1^2 + \cdots + c_k^2 = s^2 - \|(\mathrm{C} - \delta_1 \mathrm{I})^+ \mathbf{b}\|^2$$

   solves the Lagrange equations (13). Therefore the set of solutions constitutes a manifold of dimension $k - 1$.

3.3 Condition of the Secular Equation The calculation of the smallest zero λ of the secular equation (20, 21) is a delicate procedure. Even small errors $\Delta\lambda$ can result in large deviations $\Delta\mathbf{x}$

at the solution $\mathbf{x}$ and $\Delta\min$ at the minimal value min. To illustrate the point we will approximately determine the deviations $\Delta\mathbf{x}$ and $\Delta\min$ if $\Delta\lambda$ is given. We will assume that for small $\Delta\lambda$ the quantities $\Delta\mathbf{x}$ and $\Delta\min$ are essentially linearly dependent of $\Delta\lambda$.

Starting with the smallest zero λ of the explicit secular equation (20), $\mathbf{x}$ and min are computed as follows: $\mathbf{u} := (\mathrm{D} - \lambda\mathrm{I})^{-1}\mathbf{d}$, $\mathbf{z} := \mathrm{Q}\mathbf{u}$,

$$\mathbf{x} := \mathrm{P}\begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix}, \quad \text{and}$$

$$\min := \mathbf{x}^{\mathrm{T}}\mathrm{A}\mathbf{x} = \begin{bmatrix} \mathbf{y}^{\mathrm{T}} & \mathbf{z}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathrm{B} & \Gamma^{\mathrm{T}} \\ \Gamma & \mathrm{C} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix}.$$

If instead of the theoretical zero λ, the value $\lambda + \Delta\lambda$ is used the result is as follows:

$$\begin{aligned} \mathbf{u} + \Delta\mathbf{u} &= (\mathrm{D} - (\lambda + \Delta\lambda)\mathrm{I})^{-1}\,\mathbf{d} \\ &= (\mathrm{D} - \lambda\mathrm{I})^{-1}\,\mathbf{d} + (\mathrm{D} - \lambda\mathrm{I})^{-2}\,\mathbf{d}\Delta\lambda \\ &\quad + O(\Delta\lambda^2). \end{aligned}$$

From this it follows that $\Delta\mathbf{u} \approx (\mathrm{D} - \lambda\mathrm{I})^{-2}\,\mathbf{d}\Delta\lambda$. Now, we have $\mathbf{z} + \Delta\mathbf{z} = \mathrm{Q}(\mathbf{u} + \Delta\mathbf{u})$, so $\Delta\mathbf{z} = \mathrm{Q}\Delta\mathbf{u}$, and

$$\mathbf{x} + \Delta\mathbf{x} = \mathrm{P}\begin{bmatrix} \mathbf{y} \\ \mathbf{z} + \Delta\mathbf{z} \end{bmatrix}$$

implies

$$\Delta\mathbf{x} = \mathrm{P}\begin{bmatrix} 0 \\ \Delta\mathbf{z} \end{bmatrix}.$$

Finally

$$\begin{aligned} \min + \Delta\min &= (\mathbf{x} + \Delta\mathbf{x})^{\mathrm{T}}\mathrm{A}(\mathbf{x} + \Delta\mathbf{x}) \\ &= \mathbf{x}^{\mathrm{T}}\mathrm{A}\mathbf{x} + 2\mathbf{x}^{\mathrm{T}}\mathrm{A}\Delta\mathbf{x} + \Delta\mathbf{x}^{\mathrm{T}}\mathrm{A}\Delta\mathbf{x} \end{aligned}$$

so that

$$\Delta \min \approx 2\mathbf{x}^{\mathrm{T}}\mathrm{A}\Delta\mathbf{x}.$$

Now considering that

$$\begin{aligned}
2\mathbf{x}^{\mathrm{T}}\mathrm{A}\Delta\mathbf{x} &= 2\mathbf{x}^{\mathrm{T}}\mathrm{P}\mathrm{P}^{\mathrm{T}}\mathrm{A}\mathrm{P}\mathrm{P}^{\mathrm{T}}\Delta\mathbf{x} \\
&= 2\begin{bmatrix} \mathbf{y}^{\mathrm{T}} & \mathbf{z}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathrm{B} & \Gamma^{\mathrm{T}} \\ \Gamma & \mathrm{C} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \Delta\mathbf{z} \end{bmatrix} \\
&= 2\mathbf{y}^{\mathrm{T}}\Gamma^{\mathrm{T}}\Delta\mathbf{z} + 2\mathbf{z}^{\mathrm{T}}\mathrm{C}\Delta\mathbf{z} \\
&= 2(\mathbf{z}^{\mathrm{T}}\mathrm{C} - \mathbf{b}^{\mathrm{T}})\Delta\mathbf{z},
\end{aligned}$$

we get for the deviations $\Delta\mathbf{x}$ and $\Delta \min$ the values

$$\begin{aligned}
\Delta\mathbf{x} &= \mathrm{P}\begin{bmatrix} \mathbf{0} \\ \mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d} \end{bmatrix}\Delta\lambda \\
&\equiv \kappa(\mathbf{x})\Delta\lambda, \quad \text{and} \qquad (22) \\
\Delta \min &= 2(\mathbf{z}^{\mathrm{T}}\mathrm{C} - \mathbf{b}^{\mathrm{T}})\mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d}\Delta\lambda \\
&\equiv \kappa(\min)\Delta\lambda. \qquad (23)
\end{aligned}$$

Here the quantities

$$\kappa(\mathbf{x}) := \mathrm{P}\begin{bmatrix} \mathbf{0} \\ \mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d} \end{bmatrix} \qquad (24)$$

$$\kappa(\min) := 2(\mathbf{z}^{\mathrm{T}}\mathrm{C} - \mathbf{b}^{\mathrm{T}})\mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d} \qquad (25)$$

denote the *condition vector* of $\mathbf{x}$ and the *condition number* of min. In the actual computation $\kappa(\mathbf{x})$ and $\kappa(\min)$ can be calculated as well, and one can get an estimate of the numerical error.

The value of $||\kappa(\mathbf{x})||$ is bounded as follows:

$$\begin{aligned} ||\kappa(\mathbf{x})|| &= ||\mathrm{P}\begin{bmatrix} 0 \\ \mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d} \end{bmatrix}|| \\ &\le ||(\mathrm{D}-\lambda\mathrm{I})^{-2}||\,||\mathbf{d}|| = \frac{||\mathbf{d}||}{(\delta_1-\lambda)^2}. \end{aligned} \tag{26}$$

If $\mathbf{d}$ happens to be an eigenvector to the eigenvalue δ_1 of D, then we have an equality. Similarly, we can bound the condition number of min:

$$\begin{aligned} |\kappa(\min)| &= |2\mathbf{x}^{\mathrm{T}}\mathrm{AP}\begin{bmatrix} 0 \\ \mathrm{Q}(\mathrm{D}-\lambda\mathrm{I})^{-2}\mathbf{d} \end{bmatrix}| \\ &< 2||\mathrm{A}||\,||(\mathrm{D}-\lambda\mathrm{I})^{-2}||\,||\mathbf{d}|| \\ &= 2||\mathrm{A}||\frac{||\mathbf{d}||}{(\delta_1-\lambda)^2}. \end{aligned} \tag{27}$$

However, we have a strict inequality now. We can prove this by contradiction. Without loss of generality, let us assume P = I and Q = I. That is, we could apply the transformations (4) and (14) beforehand and start right away with a matrix A, where the trailing n-by-n submatrix is diagonal, and with an upper triangular matrix N. Suppose that we are given a problem where the bound on $|\kappa(\min)|$ is attained. This essentially implies, that

$$|\mathbf{x}^{\mathrm{T}}\mathrm{A}\Delta\mathbf{x}| = ||\mathbf{x}||\,||\mathrm{A}||\,||\Delta\mathbf{x}||.$$

This equality can only be satisfied if $\Delta\mathbf{x}$ is an eigenvector to the largest eigenvalue (in absolute value) of A. Furthermore, $\mathbf{x}$ and $\Delta\mathbf{x}$ must be parallel, i.e. $\mathbf{x}$ is a multiple of $\Delta\mathbf{x}$. As the first m elements of $\Delta\mathbf{x}$ are zero, so are the corresponding elements of $\mathbf{x}$. (6) implies $\mathbf{y} = \mathbf{0}$, and from (9) it follows that $\mathbf{b} = \mathbf{0}$. But from (16) we have $\mathbf{d} = \mathbf{0}$, so the desired smallest λ of the Lagrange equations (13) lies in the spectrum $\lambda(\mathrm{C})$, and we would not be solving the secular equation at all.

Thus it is obvious that we have to face large errors if the smallest zero λ of the secular equation (20, 21) is near the smallest eigenvalue δ_1 of C. Since the norm of the matrix A is normally bigger than 1 an inaccurately determined zero affects the minimal value min more than the solution vector $\mathbf{x}$.

4 ZERO FINDER Now, we want to calculate the smallest zero of the secular equation (20) with $d_k \neq 0$. We will solve it by using an iterative method. Suppose we know an approximation $\lambda^{(i)}$. Then we can approximate $f(\lambda)$ with the replacement function

$$g(\lambda) = \frac{a}{(b-\lambda)^2} - s^2 \tag{28}$$

in such a way that

$$g(\lambda^{(i)}) = f(\lambda^{(i)})$$
$$g'(\lambda^{(i)}) = f'(\lambda^{(i)}).$$

The zero of $g(\lambda)$ will determine the next approximation $\lambda^{(i+1)}$.

A short calculation yields the values

$$a = 4\frac{\left(f(\lambda^{(i)}) + s^2\right)^3}{f'^2(\lambda^{(i)})}$$
$$b = \lambda^{(i)} + 2\frac{f(\lambda^{(i)}) + s^2}{f'(\lambda^{(i)})}$$

and for the zero $\lambda^{(i+1)} = b - \sqrt{a}/s$ of the replacement function $g(\lambda)$ we get

$$\lambda^{(i+1)} = \lambda^{(i)} - 2\frac{f(\lambda^{(i)}) + s^2}{f'(\lambda^{(i)})}\left(\frac{\sqrt{f(\lambda^{(i)}) + s^2}}{s} - 1\right).$$

It can be shown that this iteration process will yield a strictly decreasing sequence of approximations $\lambda^{(i)}$. The reader is referred to [9,12].

4.1 Initial Value Now, in order to start the iteration we need to construct an initial value. For a first guess the reduced secular equation

$$\left(\frac{d_k}{\delta_k - \lambda}\right)^2 - s^2 = 0$$

is useful. This leads to the initial value

$$\lambda^{(0)} = \delta_k - \frac{|d_k|}{s}.$$

For this $\lambda^{(0)}$ it is obvious that $\lambda^{(0)} < \delta_k$ with $f(\lambda^{(0)}) \geq 0$.

4.2 Stopping Criterion In theory, the iteration process yields a strictly decreasing sequence of approximations $\lambda^{(i)}$, but this property does not persist in finite arithmetic. It is therefore reasonable to terminate the iteration when the strict monotonicity is lost, namely if $\lambda^{(i+1)} \geq \lambda^{(i)}$. This method has the advantage that it is machine-independent and that it does not need any knowledge of machine accuracy.

4.3 Implicit Secular Equation If we do not want to compute the eigenvalue decomposition (14) of C, we have to consider the evaluation of the *implicit secular function*

$$f(\lambda) = \mathbf{b}^{\mathrm{T}}(\mathrm{C} - \lambda \mathrm{I})^{-2}\mathbf{b} - s^2 \tag{29}$$

and its derivative

$$f'(\lambda) = 2\mathbf{b}^{\mathrm{T}}(\mathrm{C} - \lambda \mathrm{I})^{-3}\mathbf{b}.$$

With the definitions

$$\begin{aligned} \mathbf{u} &:= (\mathrm{C} - \lambda \mathrm{I})^{-1}\mathbf{b} \\ \mathbf{u}' &:= (\mathrm{C} - \lambda \mathrm{I})^{-1}\mathbf{u}, \end{aligned}$$

these values can be expressed as

$$\begin{aligned} f(\lambda) &= \mathbf{u}^{\mathrm{T}}\mathbf{u} - s^2 \\ f'(\lambda) &= 2\mathbf{u}^{\mathrm{T}}\mathbf{u}'. \end{aligned}$$

Therefore each iteration step requires the solution of two linear systems with the matrix $C - \lambda I$.

Again with the explicit secular equation, the quantity

$$\lambda^{(0)} = \delta_k - \frac{|d_k|}{s}$$

yields an initial value with $f(\lambda^{(0)}) \geq 0$ and $\lambda^{(0)} < \delta_k$. If the column vector $\mathbf{q}_k$ of Q were known, we could compute the value d_k, since (16) implies $d_k = \mathbf{q}_k^T \mathbf{b}$. If, however, we have an eigenvector ξ with $||\xi|| = 1$ corresponding to the smallest eigenvalue δ_1 of C, it is straightforward to use the quantity

$$\lambda^{(0)} := \delta_1 - \frac{|\xi^T \mathbf{b}|}{s}$$

as an initial value. Then it can be shown [12] that $f(\lambda^{(0)}) \geq 0$ and $\lambda^{(0)} < \delta_1$ holds when $\xi^T \mathbf{b} \neq 0$.

5 QUADRATIC EIGENVALUE PROBLEM The two previously mentioned methods have the property that they reduce the problem to finding the solution of a one-dimensional secular equation. Considering the problem from another point of view the Lagrange equations (13) can be reduced to a quadratic eigenvalue problem. For the derivation let us assume $\lambda \notin \lambda(C)$. In this case $\mathbf{z}$ can be written as

$$\mathbf{z} = (C - \lambda I)^{-1} \mathbf{b}.$$

Taking into account the normalization condition for $\mathbf{z}$ we get the secular function

$$f(\lambda) = \mathbf{b}^T (C - \lambda I)^{-2} \mathbf{b} - s^2,$$

of which the zeroes are to be computed. The task looks different if we make the definition

$$\gamma := (C - \lambda I)^{-2} \mathbf{b}$$

so that

$$(C - \lambda I)^2 \gamma = \mathbf{b}.$$

Instead of the secular equations, we have to solve the system

$$\mathbf{b}^{\mathrm{T}}\gamma - s^2 = 0 \tag{30}$$

$$(\mathrm{C} - \lambda \mathrm{I})^2 \gamma = \mathbf{b}. \tag{31}$$

The first condition (30) can also be formulated as

$$1 = \frac{1}{s^2}\mathbf{b}^{\mathrm{T}}\gamma.$$

Using this factor 1 as a coefficient of **b** in (31) we get the *quadratic eigenvalue problem*

$$(\mathrm{C} - \lambda \mathrm{I})^2 \gamma = \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma. \tag{32}$$

Note that the restriction $\lambda \notin \lambda(\mathrm{C})$ is no longer necessary. Of course, we must face the fact that the set of solutions for λ has been extended by these manipulations, for two equations cannot be formulated as a single one without consequences. Subsequently we will compare the solutions of the quadratic eigenvalue problem (32) with those of the Lagrange equations (13).

5.1 Solvability We show the following. Assume λ and $\mathbf{z}$ fulfil the Lagrange equations

$$\mathrm{C}\mathbf{z} = \lambda \mathbf{z} + \mathbf{b}$$
$$\mathbf{z}^{\mathrm{T}}\mathbf{z} = s^2.$$

THEOREM 5.1. The quadratic eigenvalue problem

$$(\mathrm{C} - \lambda \mathrm{I})^2 \gamma = \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma$$

has a solution for this λ.

PROOF. In our proof, we have to distinguish whether λ lies in the spectrum $\lambda(\mathrm{C})$ or not.

Case 1: $\lambda \in \lambda(\mathrm{C})$. Let γ be an eigenvector of C to the eigenvalue λ. Then $\mathbf{b} = (\mathrm{C}-\lambda\mathrm{I})\mathbf{z}$ implies $\mathbf{b}^{\mathrm{T}}\gamma = \mathbf{z}^{\mathrm{T}}(\mathrm{C}-\lambda\mathrm{I})\gamma = 0$.

From this it follows that

$$(\mathrm{C}-\lambda\mathrm{I})^2\gamma = \mathbf{0}$$
$$\frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma = \mathbf{0},$$

and hence γ satisfies the eigenvalue equation.

Case 2: $\lambda \notin \lambda(\mathrm{C})$. With the definition

$$\gamma := (\mathrm{C}-\lambda\mathrm{I})^{-1}\mathbf{z}$$

it follows that

$$\begin{aligned}(\mathrm{C}-\lambda\mathrm{I})^2\gamma &= (\mathrm{C}-\lambda\mathrm{I})\mathbf{z} = \mathbf{b}\\ \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma &= \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}(\mathrm{C}-\lambda\mathrm{I})^{-1}\mathbf{z}\\ &= \frac{1}{s^2}\mathbf{b}\mathbf{z}^{\mathrm{T}}\mathbf{z} = \mathbf{b},\end{aligned}$$

and again γ satisfies the eigenvalue equation.

Therefore, we can construct a solution γ of the quadratic eigenvalue problem (32) in both cases. □

Conversely, we can assume that λ and γ fulfil the quadratic eigenvalue equation

$$(\mathrm{C}-\lambda\mathrm{I})^2\gamma = \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma, \qquad \lambda \notin \lambda(\mathrm{C}).$$

THEOREM 5.2. λ and $\mathbf{z} := (\mathrm{C}-\lambda\mathrm{I})^{-1}\mathbf{b}$ fulfil the Lagrange equations

$$\mathrm{C}\mathbf{z} = \lambda\mathbf{z} + \mathbf{b}$$
$$\mathbf{z}^{\mathrm{T}}\mathbf{z} = s^2.$$

PROOF. Obviously the first equation is satisfied. Multiplying the quadratic eigenvalue equation (32) by $(\mathrm{C}-\lambda\mathrm{I})^{-2}$ we get

$$\gamma = \frac{1}{s^2}\left(\mathbf{b}^{\mathrm{T}}\gamma\right)(\mathrm{C}-\lambda\mathrm{I})^{-2}\mathbf{b} \neq \mathbf{0}.$$

This implies $\mathbf{b}^{\mathrm{T}}\gamma \neq 0$ and

$$(\mathrm{C} - \lambda\mathrm{I})^{-2}\mathbf{b} = \frac{s^2}{\mathbf{b}^{\mathrm{T}}\gamma}\gamma.$$

The square of the norm of $\mathbf{z}$ becomes:

$$\mathbf{z}^{\mathrm{T}}\mathbf{z} = \mathbf{b}^{\mathrm{T}}(\mathrm{C} - \lambda\mathrm{I})^{-2}\mathbf{b} = \mathbf{b}^{\mathrm{T}}\frac{s^2}{\mathbf{b}^{\mathrm{T}}\gamma}\gamma = s^2\frac{\mathbf{b}^{\mathrm{T}}\gamma}{\mathbf{b}^{\mathrm{T}}\gamma} = s^2,$$

with which the second equation is satisfied too. □

Finally, we assume that λ and γ fulfil the quadratic eigenvalue equation

$$(\mathrm{C} - \lambda\mathrm{I})^2\gamma = \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma, \qquad \lambda \in \lambda(\mathrm{C}).$$

We define $\mathbf{u} := (\mathrm{C} - \lambda\mathrm{I})^{+}\mathbf{b}$.

THEOREM 5.3. For the solvability of the Lagrange equations

$$\begin{aligned} \mathrm{C}\mathbf{z} &= \lambda\mathbf{z} + \mathbf{b} \\ \mathbf{z}^{\mathrm{T}}\mathbf{z} &= s^2 \end{aligned}$$

we can make the following distinction: If the Lagrange equations are inconsistent, i.e. $(\mathrm{C} - \lambda\mathrm{I})\mathbf{u} \neq \mathbf{b}$, or if $\mathbf{u}^{\mathrm{T}}\mathbf{u} > s^2$, then there is no solution for this λ. On the other hand, if the equations are consistent, we have a unique solution $\mathbf{z} = \mathbf{u}$ for $\mathbf{u}^{\mathrm{T}}\mathbf{u} = s^2$, and several solutions for $\mathbf{u}^{\mathrm{T}}\mathbf{u} < s^2$.

PROOF. If $(\mathrm{C} - \lambda\mathrm{I})\mathbf{u} = (\mathrm{C} - \lambda\mathrm{I})(\mathrm{C} - \lambda\mathrm{I})^{+}\mathbf{b} \neq \mathbf{b}$, then $\mathbf{b} \notin \mathcal{R}(\mathrm{C} - \lambda\mathrm{I})$, and the Lagrange equations have no solution for this λ. Let be $\mathbf{b} \in \mathcal{R}(\mathrm{C} - \lambda\mathrm{I})$ in the following. Then $(\mathrm{C} - \lambda\mathrm{I})\mathbf{u} = (\mathrm{C} - \lambda\mathrm{I})(\mathrm{C} - \lambda\mathrm{I})^{+}\mathbf{b} = \mathbf{b}$, and the first Lagrange equation is satisfied.

The vector $\mathbf{u}$ denotes the solution with smallest norm of the equation $\mathrm{C}\mathbf{u} = \lambda\mathbf{u} + \mathbf{b}$. Therefore, if $\|\mathbf{u}\| > s$, the normalization constraint $\mathbf{z}^{\mathrm{T}}\mathbf{z} = s^2$ cannot be satisfied, and we have no solution of the Lagrange equations for this λ. In the case of $\|\mathbf{u}\| = s$, the vector $\mathbf{u}$ denotes the unique solution.

Finally, let us assume that $\|\mathbf{u}\| < s$. The vectors $\xi^{(1)}, \ldots, \xi^{(k)}$ shall be the k orthonormal eigenvectors corresponding to the eigenvalue λ. As

$$\mathbf{u} \perp \mathcal{N}(\mathrm{C} - \lambda\mathrm{I})$$
$$\xi^{(i)} \in \mathcal{N}(\mathrm{C} - \lambda\mathrm{I}),$$

all the vectors

$$\mathbf{z} = \mathbf{u} + c_1\xi^{(1)} + \cdots + c_k\xi^{(k)}$$

with

$$c_1^2 + \cdots + c_k^2 = s^2 - ||\mathbf{u}||^2$$

are solutions of the Lagrange equations. Therefore, the set of solutions constitutes a manifold of dimension $k - 1$. □

Discussing the solvability of the Lagrange equations (13) we derived the result that the solution with the smallest λ must satisfy $\lambda \leq \delta_1$, where δ_1 denotes the smallest eigenvalue of C. Let λ be hereafter the smallest eigenvalue of the quadratic eigenvalue equation (32). Together with the forementioned theorems we can make the following distinction:

1. It can hold that $\lambda < \delta_1$. Then λ lies outside the spectrum $\lambda(\mathrm{C})$, and $\mathbf{z} := (\mathrm{C} - \lambda\mathrm{I})^{-1}\mathbf{b}$ fulfils the Lagrange equations (13). The solution is unique.
2. Now let $\lambda = \delta_1$. The vector $\mathbf{u} := (\mathrm{C} - \lambda\mathrm{I})^{+}\mathbf{b}$ must satisfy the equation $\mathrm{C}\mathbf{u} = \lambda\mathbf{u} + \mathbf{b}$. If $\mathbf{u}^{\mathrm{T}}\mathbf{u} = s^2$, then $\mathbf{z} := \mathbf{u}$ is the unique solution of the Lagrange equations (13).
3. Finally, if $\mathbf{u}^{\mathrm{T}}\mathbf{u} < s^2$, we must find an eigenvector ξ to the eigenvalue λ of C with $\xi^{\mathrm{T}}\xi = s^2 - \mathbf{u}^{\mathrm{T}}\mathbf{u}$. Then, $\mathbf{z} := \mathbf{u} + \xi$ represents one of the many solutions of the Lagrange equations (13).

5.2 Solving the Quadratic Eigenvalue Problem The quadratic eigenvalue problem (32) can be reduced to an ordinary eigenvalue problem by properly chosen transformations. With the definition

$$\eta := (\mathrm{C} - \lambda\mathrm{I})\gamma \tag{33}$$

the following equations can be established:

$$C\gamma - \eta = \lambda\gamma$$
$$C\eta - \frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}}\gamma = \lambda\eta.$$

In matrix terms this leads to:

$$\begin{bmatrix} C & -I \\ -\frac{1}{s^2}\mathbf{b}\mathbf{b}^{\mathrm{T}} & C \end{bmatrix} \begin{bmatrix} \gamma \\ \eta \end{bmatrix} = \lambda \begin{bmatrix} \gamma \\ \eta \end{bmatrix}. \tag{34}$$

Thus we have transformed the original quadratic eigenvalue problem into an equivalent linear one, that can be solved with traditional methods.

6 NUMERICAL RESULTS

Finally we present some results of the numerical experiments. All calculations were carried out on a VAX 8600 with Floating Point Accelerator (single precision, 32 bit reals) under VMS 4.6. The IMSL-Library Edition 10.0 / Version 1.0 served for the basic computations. The input data was produced by a random number generator. As a reference solution we employed an implementation in double precision. The tables 1 to 3 summarize the results. The individual columns have the following meaning:

λ : smallest zero of the secular equation
min : minimum
Norm: norm of the solution $\mathbf{x}$
Error: norm of the difference $\mathbf{x} - \mathbf{x}_{\text{Reference Solution}}$
Time: CPU-time

Furthermore the caption contains the dimension n of the problem together with the condition numbers (24, 25) of the solutions.

Table 1: $n = 10$ $\|\kappa(\mathrm{x})\| = 0.104$ $\kappa(\min) = -0.282$

Algorithm	λ	min	Norm	Error	Time
Reference	-4.220885	-0.677439	1.000000		
Expl. Sec. Eq.	-4.220886	-0.677439	1.000000	$2.0 \cdot 10^{-7}$	0.05 sec
Impl. Sec. Eq.	-4.220886	-0.677439	1.000000	$6.5 \cdot 10^{-8}$	0.10 sec
Quadr. EVP	-4.220890	-0.677438	1.000000	$5.4 \cdot 10^{-7}$	0.23 sec

Table 2: $n = 45$ $\|\kappa(\mathrm{x})\| = 930.8$ $\kappa(\min) = -13131$

Algorithm	λ	min	Norm	Error	Time
Reference	-7.054650	-7.054494	1.000000		
Expl. Sec. Eq.	-7.054657	-7.052025	0.999825	$1.8 \cdot 10^{-4}$	1.16 sec
Impl. Sec. Eq.	-7.054650	-7.060933	1.000456	$4.6 \cdot 10^{-4}$	2.60 sec
Quadr. EVP	-7.055462	-2.290894	0.569900	$4.3 \cdot 10^{-1}$	10.4 sec

Based on the numerical calculations the three solutions can be judged as follows:

Explicit secular equation. The zero of the secular equation (20) is determined to machine precision. The accuracy that can be expected by the condition numbers (24, 25) is achieved.

Implicit secular equation. This method achieves the same accuracy as the first one. However the calculation of an eigenvalue decomposition (14) is replaced by the determination of a generalized inverse which is in no way cheaper than the former operation.

Quadratic eigenvalue problem. It turns out that the smallest eigenvalue of the general matrix (34) can be calculated only very inexactly. We suppose that the transformation into a quadratic eigenvalue problem (32) impairs the condition of the problem. With large condition numbers (24, 25) all decimal places can be incorrect.

Table 3: $n = 100$ $||\kappa(\mathbf{x})|| = 6.019$ $\kappa(\min) = -125.2$

Algorithm	λ	min	Norm	Error	Time
Reference	-11.48939	-11.07225	1.000000		
Expl. Sec. Eq.	-11.48939	-11.07225	1.000000	$5.3 \cdot 10^{-6}$	9.90 sec
Impl. Sec. Eq.	-11.48939	-11.07230	1.000002	$2.2 \cdot 10^{-6}$	24.5 sec
Quadr. EVP	-11.48976	-11.02595	0.997983	$2.2 \cdot 10^{-3}$	104. sec

ACKNOWLEDGEMENT We wish to thank Professor A. M. Lesk who stimulated the research described in this paper through a personal communication.

REFERENCES

[1] N. R. DRAPER, *"Ridge Analysis" of Response Surfaces*, Technometrics, 5 (1963), pp. 469–479.

[2] G. E. FORSYTHE AND G. H. GOLUB, *On the stationary values of a second-degree polynomial on the unit sphere*, SIAM J. Appl. Math., 13 (1965), pp. 1050–1068.

[3] W. GANDER, *Least Squares with a Quadratic Constraint*, Numer. Math., 36 (1981), pp. 291–307.

[4] G. H. GOLUB, *Some Modified Matrix Eigenvalue Problems*, SIAM Review, 15 (1973), pp. 318–334.

[5] G. H. GOLUB, M. HEATH AND G. WAHBA, *Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter*, Technometrics, 21 (1979), pp. 215–223.

[6] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1983.

[7] J. J. MORÉ, *The Levenberg-Marquardt Algorithm: Implementation and Theory*, in Proc. of the Biennial Conference held at Dundee, ed. A. Dold and B. Eckmann, Springer-Verlag, 1978, pp. 105–116.

[8] J. J. MORÉ AND D. C. SORENSEN, *Computing a Trust Region Step*, SIAM J. Sci. Stat. Comput., 4 (1983), pp. 553–572.

[9] CHR. H. REINSCH, *Smoothing by Spline Functions. II*, Numer. Math., 16 (1971), pp. 451–454.

[10] E. SPJØTVOLL, *A Note on a Theorem of Forsythe and Golub*, SIAM J. Appl. Math., 23 (1972), pp. 307–311.

[11] E. SPJØTVOLL, *Multiple Comparison of Regression Functions*, Ann. Math. Statist., 43 (1972), pp. 1076–1088.

[12] U. VON MATT, *A Constrained Eigenvalue Problem*, Diploma Thesis, Abteilung für Informatik, ETH Zürich, 1988.