Chapter 13
Eigenvalue Problems

Abstract In this chapter, we consider various numerical methods for solving the eigenvalue problem and the application of hierarchical matrix techniques. After the introduction in Section 13.1, we discuss the LR and QR methods in Section 13.2, the vector iteration and Krylov methods in Section 13.3, and the preconditioned inverse iteration in Section 13.4. Furthermore, we consider the bisection method in Section 13.5 and the divide-and-conquer method in Section 13.6. Quite another approach is the computation of an eigenvalue distribution in Section 13.7. Finally, spectral projections are studied in Section 13.8.

13.1 Introduction

There are different kinds of eigenvalue problems depending on the properties of the matrices. The most favourable case is the *symmetric* eigenvalue problem

$$Ax = \lambda x, \qquad A = A^{\mathsf{T}} \in \mathbb{R}^{I\times I}. \tag{13.1}$$

Here it is known that all eigenvalues and eigenvectors are real and that the eigenvectors form an orthogonal basis of $\mathbb{R}^I$ (cf. Parlett [213]).

In the nonsymmetric case, eigenvalues and eigenvectors may become complex. For each eigenvalue there is at least one eigenvector, but the number of linearly independent eigenvectors may be less than $\#I$.

The generalised eigenvalue problem is

$$Ax = \lambda Bx, \tag{13.2}$$

in which the easiest case requires that A be symmetric and B positive definite.

Further generalisations lead to quadratic or nonlinear eigenvalue problems (cf. Börm–Mehl [59, §8]).

W. Hackbusch, *Hierarchical Matrices: Algorithms and Analysis*, Springer Series in Computational Mathematics 49, DOI 10.1007/978-3-662-47324-5_13

The 'eigenvalue problem' branches into a number of different questions. Possible tasks are:

T1: Determine one or few eigenvalues.

T2: Determine one or few eigenvectors.

T3: Determine all eigenvalues.

T4: Determine all eigenvectors.

T5: Determine a rough distribution of the spectrum.

The tasks **T1** and **T2** (or **T3** and **T4**) can be combined. In the case of **T5**, one is not interested in the exact location of all eigenvalues, but in a density function (details in §13.7).

Correspondingly, there are quite different approaches to the eigenvalue problem. The numerical methods make use of the following techniques:

1. similarity transformation $M \mapsto TMT^{-1}$ (cf. §13.2),
2. shift techniques $M \mapsto M - \mu I$, or generalised shifts $M \mapsto (M - \mu I)^2$, etc. (cf. §13.2),
3. vector iteration $x \mapsto Mx$ (cf. §§13.3, 13.4),
4. Krylov techniques (cf. §13.3),
5. other subspace techniques (cf. §13.9),
6. eigenvalue computation via the Rayleigh quotient (cf. §13.3),
7. bisection techniques for the search of eigenvalues (cf. §13.5),
8. spectral projections (cf. §13.8),
9. other techniques (cf. §§13.6, 13.7).

Remark 13.1. Before entering into details, we can give some general statements concerning the use of hierarchical matrices.

(a) Any iterative method using similarity transformations (item 1) is in trouble, if we perform inexact multiplications by T and T^{-1}.

(b) A shift (items 1,2) can be performed exactly, since only diagonal blocks are changed. These blocks are exactly represented by full matrices.

(c) The matrix-vector multiplication in item 4 is performed exactly.

(d) Item 6 is a generalisation of item 4, and the same comment holds. The Rayleigh quotient

$$\mu(x) := \frac{\langle Mx, x\rangle}{\langle x, x\rangle}$$

is performed exactly (use Exercise 7.2).

Large parts of this chapter summarise the results of the doctoral thesis by Mach [206] to which we refer for more details and numerical results.

13.2 LR and QR Methods

The methods in this group are iterative and are based on a suitable factorisation

$$M_i = A_i B_i \qquad (A_i \text{ regular}).$$

The next iterate

$$M_{i+1} = B_i A_i$$

can be written as $M_{i+1} = A_i^{-1} M_i A_i$, hence M_i and M_{i+1} are similar matrices; i.e., they possess the same spectrum.[1] The starting value $M_0 = M$ is the original matrix. One expects that M_i converges to a diagonal matrix (or block diagonal matrix in the case of multiple eigenvalues).

The historically older choice is the LU decomposition (cf. Rutishauser [223]), which in this context is called the LR decomposition. Hence, the LR method is the iteration

$$M_i = L_i R_i, \quad M_{i+1} = R_i L_i.$$

In the general case, L_i is a normed left triangular matrix, R_i a right triangular matrix. Then one must require that the decomposition $M_i = L_i R_i$ exist for all i. In the symmetric case we exploit that the eigenvalue problems for M_i and for the shifted matrix $M_i + \mu I$ are essentially equal, so that without loss of generality M_i can be assumed to be positive definite. Then the Cholesky decomposition with $R_i = L_i^T$ is used.

A more stable version is the QR method ([59, §5])

$$M_i = Q_i R_i, \quad M_{i+1} = R_i Q_i,$$

where Q_i is unitary and R_i as above. The iterations are accelerated by a suitable shift technique ($M_i - \mu I = L_i R_i$, $M_{i+1} = R_i L_i + \mu I$; cf. [59, §5.3]) and a deflation technique (cf. [59, §5.4]).

In principle, one can apply these iterations to any matrix M. In practice, one first applies a similarity transformation which maps M into a tridiagonal[2] matrix (cost: $\mathcal{O}(n^3)$), since the LR and QR iterations preserve this structure and are much cheaper to perform.

The crucial question is whether there is an $\mathcal{H}$-matrix format that remains unchanged under similarity transformations. In the general case, the answer is negative: the iteration $M_{i+1} = L_i^{-1} M_i L_i$ increases the local rank until all block have maximal rank. On the other hand, truncations to a fixed format might be tolerated for a fixed number of operations, but their influence is hard to estimate for an infinite process (note that the iteration is not of the fixed-point type analysed in §15.3.2).

However, there is a positive result for the special $\mathcal{H}$-matrix block partition P from §3.1 (see (3.2) and Fig. 3.1, left). Note that $P = P_{\text{weak}}$ is also mentioned in §9.3. But differently from the setting in §3.1, we use a rank distribution $r_{\text{lin}} : P \to \mathbb{N}$

[1] Another iterative method based on similarity transformations is the Jacobi method (cf. [59, §3]).

[2] This only holds for the symmetric case (13.1). In the general case, the target matrix is of Hessenberg form.

increasing linearly with the level-number,[3]

$$r_{\mathrm{lin}}(b) = \ell r \qquad \text{for } b \in P^{(\ell)} := P \cap T^{(\ell)}(I \times I), \tag{13.3}$$

using the decomposition of the block cluster tree $T(I \times I) = \bigcup_\ell T^{(\ell)}(I \times I)$ with respect to the level (cf. (A.2)). Note that this format uses $n_{\min} = 1$.

The following result is stated in Benner–Mach [36, §3.4]. This paper also contains numerical computations and comparisons.

Theorem 13.2. *If the positive definite matrix M belongs to $\mathcal{H}(r_{\mathrm{lin}}, P_{\mathrm{weak}})$ with r_{lin} and P_{weak} defined as above, all iterates stay in this format.*

Remark 13.3. Let $\#I =: n = 2^L$. The storage cost of a matrix from $\mathcal{H}(r_{\mathrm{lin}}, P_{\mathrm{weak}})$ is equal to

$$S = n \sum_{\ell=1}^{L} \min\{2^{L-\ell}, \ell r\} \le \frac{nr}{2} \left[\left(\log_2\left(\tfrac{n}{r}\right)\right)^2 + \mathcal{O}\left(\log_2\left(\tfrac{n}{r}\right)\right) \right].$$

Here we exploit the symmetric structure; i.e., only blocks of the upper triangle have to be stored.

Proof. At the level ℓ, we have $2^{\ell-1}$ blocks represented by 2κ vectors of length $2^{L-\ell}$, where $\kappa = \min\{2^{L-\ell}, \ell r\}$ is the rank. Hence, the required storage cost is $n \sum_{\ell=1}^{L} \min\{2^{L-\ell}, \ell r\}$. Let $\ell_0 \in \mathbb{N}$ be the number with $\ell_0 r \le 2^{L-\ell_0}$ and $(\ell_0 + 1)\, r > 2^{L-\ell_0-1}$. We conclude that

$$\begin{aligned} S &= n \sum_{\ell=1}^{L} \min\{2^{L-\ell}, \ell r\} = n \left(\sum_{\ell=1}^{\ell_0} \ell r + \sum_{\ell=\ell_0+1}^{L} 2^{L-\ell} \right) \\ &= n \left[\frac{1}{2} r \ell_0 \left(\ell_0 + 1\right) + 2^{L-\ell_0} - 1 \right]. \end{aligned}$$

Let $\lambda \in \mathbb{R}$ be the solution of $\lambda r = 2^{L-\lambda}$. Since $\ell_0 \le \lambda < \ell_0 + 1$, we estimate by

$$S \le n \left[\frac{1}{2} r \lambda \left(\lambda + 1\right) + 2^{L-\lambda-1} - 1 \right] = n \left[\frac{1}{2} r \lambda \left(\lambda + 2\right) - 1 \right].$$

The solution of the nonlinear equation $\lambda r = 2^{L-\lambda}$ has asymptotic behaviour $\lambda = \log_2(n/r) + \mathcal{O}(\log_2 \log_2(n/r))$ as $n/r \to \infty$. □

For instance, for $n = 2^{15} = 32\,768$ and $r = 40$, we have $\ell_0 = 6$, so that $S = \frac{1351}{40} nr = 33.8\ldots \cdot nr < 46.8\ldots \cdot nr = \frac{nr}{2} \left(\log_2(n/r)\right)^2$.

The paper [36, §3.4] also discusses the application of the LR method to matrices of the semiseparable format (cf. §3.9 and Li–Gu–Cheng [198]).

[3] Note that a similar structure also appears in Lemma 9.14 and Conclusion 9.15.

13.3 Vector Iteration and Krylov Methods

The vector iteration (or 'power iteration', cf. [59, §4]) is a method for task T3, since it computes the eigenvector corresponding to the largest eigenvalue in modulus. The iteration

$$x^{(m)} \mapsto x^{(m+1)} := \frac{1}{\|Mx^{(m)}\|} Mx^{(m)}$$

only requires matrix-vector multiplication, which can be performed exactly for hierarchical matrices M. Block versions can be used to determine the eigenvectors of the k largest eigenvalues ('simultaneous iteration', cf. [59, §4.7]).

Although the iteration yields the approximate eigenvector x, the corresponding eigenvalue is an immediate result of the Rayleigh quotient

$$\mu(x) := \langle Mx, x\rangle / \langle x, x\rangle . \tag{13.4}$$

In the symmetric case, the error of the approximate eigenvalue $\mu(x)$ is proportional to the square of the error of the eigenvector approximation x (cf. [59, Thm. 4.6]).

The Krylov method is based on the Krylov space spanned by the vectors $x^{(0)}$, $x^{(1)}, \ldots, x^{(k)}$ (cf. [59, §7]). Since nothing other than matrix-vector multiplication is required, Krylov methods apply perfectly to hierarchical matrices, exploiting the fast matrix-vector multiplication in the case of non-sparse matrices. For a convergence analysis, see Kandler–Schröder [165].

Often one needs a rough approximation of the spectral norm $\|M\|_2$. This can be achieved by vector iteration, independently of the distribution of the eigenvalues. Note that this statement is not true for approximating the corresponding eigenvector. In the latter case, the convergence speed depends on the size of the gap between the largest eigenvalues.

Theorem 13.4. *Apply the vector iteration to $M^{\mathsf{T}}M$ or MM^{T} with a random start vector $x^{(0)}$. Then the spectral norm $\|M\|_2 = \rho(M^{\mathsf{T}}M)^{1/2} = \rho(MM^{\mathsf{T}})^{1/2}$ can be determined up to a relative error ε by*

$$O\left(\frac{1}{\varepsilon}\left(\log(\operatorname{rank}(M)) + \log(1/\varepsilon)\right)\right)$$

steps of the vector iteration.

A proof can be found in Grasedyck [102, Satz 4.31]. The random nature of $x^{(0)}$ is used to ensure that the component α_1 in $x^{(0)} = \sum_{i=1}^{\operatorname{rank}(M)} \alpha_i e_i$ (e_i are the orthonormal eigenvectors of $M^{\mathsf{T}}M$ with respect to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots$) is not vanishing and has the expectation value $\alpha_1^2 / \sum_{i=1}^{\operatorname{rank}(M)} \alpha_i^2 = 1/\operatorname{rank}(M)$.

13.4 Preconditioned Inverse Iteration

The largest eigenvalues targeted in §13.3 are of interest for operators of negative order, i.e., integral operators. In the case of discretised elliptic differential equations, one looks for the smallest ones. The smallest eigenvalues can be obtained by the vector iteration applied to M^{-1} instead of M. The latter iteration is called the inverse iteration (cf. [59, §4.4]). More generally, eigenvalues close to $\mu \in \mathbb{R}$ can be obtained by the inverse iteration applied to $M - \mu I$. Let T be any regular matrix ('preconditioner'). Implicitly, the following inverse iteration

$$x^{(m+1)} = x^{(m)} - T^{-1}\left[Mx^{(m)} - \mu(x^{(m)})x^{(m)}\right] \tag{13.5}$$

for the eigenvector $\lim x^{(m)}$ is also an iteration for the eigenvalue $\lim \mu(x^{(m)})$. For $T = I$, it describes the gradient method for minimising the Rayleigh quotient (13.4). Since, in general, this iteration is rather slow, one uses an approximation $T \approx M$ which is easy to invert. Estimates about the convergence of (13.5) are given in Benner–Mach [37, Theorem 2.1].

Given some hierarchical matrix $M = M^{\mathsf{T}}$, we may compute the $\mathcal{H}$-Cholesky decomposition $M \approx LL^{\mathsf{T}} =: T$. Depending on the chosen ranks, $T^{-1} = L^{-\mathsf{T}}L^{-1}$ is a good approximation of the inverse of M. Generalising the algorithm to the simultaneous computation of the smallest d eigenvalues and corresponding eigenvectors is described in [37, Alg. 4]. Also the application to the generalised eigenvalue problem (13.2) is discussed in [37, Remark 4.1].

In order to compute eigenvectors corresponding to interior eigenvalues close to μ, one may apply these methods to $(M - \mu I)^2$ instead of $M - \mu I$ (details and numerical examples in [37, §5]).

13.5 Bisection Method

Methods for finding roots of polynomials have long been known. One approach is Sturm's chain (cf. [59, §6.1]). It allows us to determine the number of zeros in an interval $[a, b)$. Starting with global bounds $[A, B)$ of the zeros, one bisects the interval, checks $[A, (A + B)/2)$ and $[(A + B)/2, B)$, and continues recursively until one root is included in an interval $[\xi, \xi + \varepsilon)$ of sufficiently small length ε.

In the case of tridiagonal matrices, the characteristic polynomials of the principal submatrices form a Sturm chain (cf. [59, Thm. 6.4]), so that the eigenvalues of (13.1) can be determined by bisection.

Note that this methods allows us to determine any desired eigenvalue. It is not restricted to the largest or smallest part of the spectrum. Furthermore, the bisection process may be directed to one particular eigenvalue as well as to all eigenvalues. In the case that several eigenvalues are sought, the algorithm ends up with several

disjoint intervals. Therefore the bisection process can be performed in parallel for each interval.

Another indicator of the number of eigenvalues in some interval can be based on Sylvester's inertia law. The matrix M is assumed to be a symmetric to ensure a real spectrum. Matrices M' and M'' are congruent if there is a congruence transformation $M'' = TM'T^{\mathsf{T}}$ (T regular). Each matrix $M \in \mathbb{R}^{I\times I}$ is congruent to a matrix

$$\operatorname{diag}\{-1,\dots,-1,0,\dots,0,1,\dots,1\}$$

containing ν times the entry -1, ξ times 0, and $\#I - \nu - \xi$ times $+1$. The triple $(\nu, \xi, \#I - \nu - \xi)$ is called the *inertia* of M. The number of eigenvalues of M in $(-\infty, \mu)$ is equal to $\nu(\mu)$, where $\nu(\mu)$ is defined by the inertia of the shifted matrix $M - \mu I$. Since we know that $\nu(\mu) = \#I$ for $\mu > \|M\|$ and $\nu(\mu) = 0$ for $\mu < -\|M\|$, the function $\nu(\mu)$ can be used to control the bisection process (also named 'slicing the spectrum'; cf. [213, p. 51]).

Given M, the LDL decomposition

$$M - \mu I = LDL^{\mathsf{T}} \qquad (L \text{ lower triangular, } D \text{ diagonal})$$

is a particular congruence transformation. Since the number of negative diagonal entries D_{ii} is equal to $\nu(\mu)$, the LDL decomposition is the decisive tool for the bisection method.

The LDL decomposition of hierarchical matrices is first described by Lintner [202] (see also [201]). Benner–Mach [35] apply the LDL decomposition to matrices of the format $\mathcal{H}(r, P_{\mathrm{weak}})$, where the matrix blocks have constant local rank r, and prove the following result.

Remark 13.5. Application of the LDL decomposition to $M \in \mathcal{H}(r, P_{\mathrm{weak}})$ without truncation yields the triangular factor $L \in \mathcal{H}(r_{\mathrm{lin}}, P_{\mathrm{weak}})$ with r_{lin} defined in (13.3). The computational work is estimated in [35, Lemma 2.4] by[4] $O(nk^2 \log^4 n)$, where $n = \#I$.

Since the remark describes the cost for one LDL decomposition, the cost for approximating *all* eigenvalues is proportional to n^2. The number of decompositions per eigenvalue increases proportional to $\log(1/\varepsilon)$ as the required accuracy ε tends to zero.

The use of the bisection method based on an inexact LDL decomposition of *general* hierarchical matrices is possible, but because of truncation errors, the computed diagonal matrix D need not be congruent to $M - \mu I$. These problems and further details are explained in [35].

In the case of $\mathcal{H}^2$-matrices, efficient LDL decompositions and an efficient bisection method are described by Benner–Börm–Mach–Reimer [33]. These authors treat also the generalised eigenvalue problem (13.2).

[4] We expect that a proof using the true rank $\kappa = \min\{2^{L-\ell}, \ell k\}$ instead of the bound ℓk will yield an estimate with k^2 replaced by k.

13.6 Divide-and-Conquer Method

The following method described by Gördes [97] is a generalisation of the divide-and-conquer methods of Cuppen [74] and Dongarra–Sorensen [83].

In the trivial case of $M = \begin{bmatrix} M' & 0 \\ 0 & M'' \end{bmatrix}$, we can divide the eigenvalue problem for M into two problems for M' and M''. Now we consider matrices of the hierarchical format $\mathcal{H}_p$ introduced in §3.1. In the symmetric case, a matrix $M \in \mathcal{H}_p$ ($p \geq 1$) has the form

$$M = \begin{bmatrix} M' & ab^{\mathsf{T}} \\ ba^{\mathsf{T}} & M'' \end{bmatrix} \qquad \text{with } M', M'' \in \mathcal{H}_{p-1},\ \ a, b \in \mathbb{R}^{2^{p-1}}.$$

Assume that the eigenvalue problems for M' and M'' are already solved:

$$M' = Q'^{\mathsf{T}} D' Q' \text{ and } M'' = Q''^{\mathsf{T}} D'' Q''.$$

A similarity transformation by $\begin{bmatrix} Q' & 0 \\ 0 & Q'' \end{bmatrix}$ maps M into

$$\hat{M} = \begin{bmatrix} D' & \hat{a}\hat{b}^{\mathsf{T}} \\ \hat{b}\hat{a}^{\mathsf{T}} & D'' \end{bmatrix}, \quad \hat{a} := Q'a\,, \quad \hat{b} := Q''b\,.$$

We define the rational function

$$r(\lambda) := \left(\sum_{i=1}^{2^{p-1}} \frac{a_i^2}{d_i' - \lambda} \right) \left(\sum_{i=1}^{2^{p-1}} \frac{b_i^2}{d_i'' - \lambda} \right) - 1\,.$$

Then the following identity holds (cf. [97, Satz 4]):

$$\det(\hat{M} - \lambda I) = -r(\lambda) \cdot \det(D' - \lambda I) \cdot \det(D'' - \lambda I)\,.$$

In the standard case, the zeros of $\det(D' - \lambda I) \cdot \det(D'' - \lambda I)$ and the poles of $r(\lambda)$ cancel, so that the zeros of r describe all eigenvalues of $\hat{M}$. Exceptions are:
(a) $a_i = 0$ for some i,
(b) $b_i = 0$ for some i,
(c) D' has multiple eigenvalues d_i',
(d) D'' has multiple eigenvalues d_i''. In these cases, either d_i' and/or d_i'' are an additional eigenvalue of $\hat{M}$.

The zeros of $r(\cdot)$ can be determined by a combination of Newton's method and bisection. Having determined the eigenvalue, we can easily obtain the eigenvectors by the inverse iteration (cf. §13.4).

13.7 Eigenvalue Distribution

The following algorithm corresponds to the method of Haydock–Heine–Kelly [153]. This and further methods are discussed by Lin–Saad–Yang [200].

We assume a real spectrum with eigenvalues $(\lambda_i)_{i\in I}$ (multiple eigenvalues included). Since $\#I$ is expected to be very large, often the tuple $(\lambda_i)_{i\in I}$ is not the appropriate information. Instead one likes to visualise the distribution by the graph of the function

$$\Lambda(\xi) = \sum_{\nu\in I} \phi(\xi - \lambda_\nu)$$

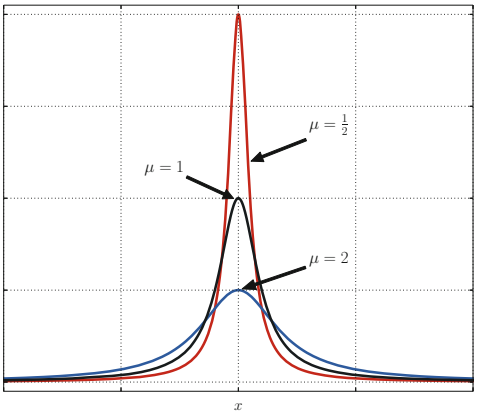

Fig. 13.1 Function ϕ_μ for $\mu = \frac{1}{2}, 1, 2$

with a suitable, singly peaked bump function ϕ. Besides $\phi \geq 0$, we may require $\int_{\mathbb{R}} \phi(\xi)\mathrm{d}\xi = 1$ so that

$$\int_{\mathbb{R}} \Lambda(\xi)\mathrm{d}\xi = n := \#I.$$

A typical example for ϕ is

$$\phi_\mu(\xi) = \frac{1}{\pi}\frac{\mu}{\xi^2 + \mu^2} \qquad (\mu > 0). \tag{13.6}$$

Depending on the application field, ϕ_μ is called the Cauchy distribution or the Lorentzian function (cf. Fig. 13.1). For $\mu \searrow 0$, the function ϕ_μ approaches the delta distribution. Note that $\Lambda(\xi)$ can be interpreted as the convolution of ϕ_μ with the (exact) spectral distribution

$$\lambda(t) := \sum_{\nu\in I} \delta(t - \lambda_\nu),$$

i.e., $\Lambda = \lambda \star \phi$ is a smoothed version of λ.

It will turn out that the distribution $\Lambda(\xi)$ with $\phi = \phi_\mu$ from (13.6) can be obtained without solving for all eigenvalues explicitly. We consider M with a complex shift $\xi + \mathrm{i}\mu$ ($\xi, \mu \in \mathbb{R}$). The determinant of $M - (\xi + \mathrm{i}\mu)I$ is the product $\prod_{\nu\in I}(\lambda_\nu - (\xi + \mathrm{i}\mu))$ of its eigenvalues. The logarithm of the absolute value is

$$\begin{aligned}\log|\det(M - (\xi + \mathrm{i}\mu)I)| &= \log\left|\prod_{\nu\in I}(\lambda_\nu - (\xi + \mathrm{i}\mu))\right| = \sum_{\nu\in I}\log|\lambda_\nu - (\xi + \mathrm{i}\mu)| \\ &= \sum_{\nu\in I}\frac{1}{2}\log\left[(\lambda_\nu - \xi)^2 + \mu^2\right].\end{aligned}$$

The derivative with respect to μ yields $\Lambda = \lambda \star \phi_\mu$ with an additional factor π:

$$\frac{\mathrm{d}}{\mathrm{d}\mu} \log |\det (M - (\xi + \mathrm{i}\mu)I)| = \sum_{\nu \in I} \frac{\mu}{(\lambda_\nu - \xi)^2 + \mu^2} = \pi \lambda \star \phi_\mu \,.$$

For the numerical realisation, we use an LU decomposition of the shifted matrix:

$$M - (\xi + \mathrm{i}\mu)I = LU.$$

Since $\det(L) = 1$, we have $\det (M - (\xi + \mathrm{i}\mu)I) = \det(U) = \prod_{\nu \in I} U_{\nu\nu}$ and

$$L(\mu) := \log |\det (M - (\xi + \mathrm{i}\mu)I)| = \sum_{\nu \in I} \log |U_{\nu\nu}| \,.$$

Instead of the derivative, we use the difference quotient

$$\frac{L(\mu + h\mu) - L(\mu - h\mu)}{2h\mu} = \frac{\mathrm{d}L}{\mathrm{d}\mu}(\vartheta\mu) = \sum_{\nu \in I} \frac{\vartheta\mu}{(\lambda_\nu - \xi)^2 + \vartheta^2\mu^2},$$

where $\vartheta = \vartheta_\xi \in (1 - h, 1 + h)\,,\ h \in (0, 1),$ is some intermediate value.

We conclude that each evaluation of the function $\Lambda(\xi)$ requires two $\mathcal{H}$-LU decompositions. All LU decompositions are completely independent and can be computed in parallel.

A matrix representation of the true derivative is the imaginary part of the trace of the inverse:

$$\Im\mathfrak{m}\,\mathrm{trace}\big((M - (\xi + \mathrm{i}\mu)I)^{-1} \big) = \sum_{\nu \in I} \frac{\mu}{(\lambda_\nu - \xi)^2 + \mu^2}.$$

However, the hierarchical inverse is more costly than the $\mathcal{H}$-LU decompositions.

The choice of μ depends on the (average) gaps between the eigenvalues. The gap is equal to one for the uniform distribution $\lambda_\nu = \nu \in \mathbb{Z}$. In this case, the choice $\mu = 1$ is satisfactory since $\Lambda(\xi)$ is almost a constant. The maximal values at $\xi \in \mathbb{Z}$ and minimal value at $\xi \in 1/2 + \mathbb{Z}$ differ by less than 1 per cent.

Examples of spectral distributions appearing in computational biology can be found in Banerjee–Jost [9]. Here the eigenvalues belong to the graph Laplacian, where the size of the graph may be very large.

Assume that the eigenvalues are in $[0, 2]$ (as in the cases studied in [9]). In the case of $n = 200$ eigenvalues, a graph of $\Lambda(\cdot)$ with the step size $\mu = 0.01$ would be sufficient. The resolution by $\mu = 0.01$ need not be changed if n becomes much larger (unless we want to see a graph of $\Lambda(\cdot)$, e.g., in $[0.92, 0.93]$). Hence, the plot of the function $\Lambda(\cdot)$ requires computational work almost proportional to n.

13.8 Spectral Projections

In §14.1.3, we shall discuss Cauchy's integral $\frac{1}{2\pi\mathrm{i}} \oint_{\mathcal{C}} (\zeta I - M)^{-1} f(\zeta)\,\mathrm{d}\zeta$ for certain functions f. Here, the integration curve $\mathcal{C}$ may be regarded as the boundary of a bounded, complex domain $\Omega \subset \mathbb{C}$. In the particular case of $f(\zeta) = \zeta$ and $\sigma(M) \subset \Omega$, the residue theorem states the identity $M = \frac{1}{2\pi\mathrm{i}} \oint_{\mathcal{C}} (\zeta I - M)^{-1} \zeta\,\mathrm{d}\zeta$.

Next, we consider a subset $\sigma' \subset \sigma(M)$ of the spectrum and construct Ω such that

$$\sigma' = \sigma(M) \cap \Omega, \qquad \sigma(M) \cap \partial\Omega = \emptyset.$$

Then the integral along $\mathcal{C} := \partial\Omega$ yields the matrix

$$M_{\sigma'} := \frac{1}{2\pi\mathrm{i}} \oint_{\mathcal{C}} (\zeta I - M)^{-1} \zeta \mathrm{d}\zeta.$$

$M_{\sigma'}$ commutes with M. In the diagonalisable case of $M = T^{-1} D\, T$ with $D = \mathrm{diag}\{\lambda_i : i \in I\}$, we have

$$M_{\sigma'} = T^{-1} D_{\sigma'} T \quad \text{with } D_{\sigma'} = \mathrm{diag}\{\lambda'_i : i \in I\}, \;\; \lambda'_i := \begin{cases} \lambda_i & \text{if } \lambda_i \in \sigma', \\ 0 & \text{if } \lambda_i \notin \sigma'. \end{cases}$$

The integral can be approximated by a quadrature rule. Quadrature points at $\{\zeta_\nu : 1 \le \nu \le N\} \subset \mathcal{C}$ lead to an expression of the form

$$M_{\sigma'} \approx \sum_{\nu=1}^{N} \omega_\nu \left(\zeta_\nu I - M\right)^{-1}.$$

In principle, one can approximate the resolvents $(\zeta_\nu I - M)^{-1}$ by hierarchical matrices. However, it is cheaper to use the LU decomposition $\zeta_\nu I - M \approx L_\nu U_\nu$.

Choosing subsets σ' with moderate cardinality $\#\sigma'$, we can apply traditional methods, e.g., Krylov based methods to solve the complete eigenvalue problem for $M_{\sigma'}$. Note that the Krylov method only requires matrix-vector multiplications, which here take the form $x \mapsto U_\nu^{-1} L_\nu^{-1} x$.

Another approach is used in Hackbusch–Kress [148]. Assume a real spectrum $\sigma(M) \subset \mathbb{R}$ with eigenvalues λ_i ordered by size. Let a and b be numbers with

$$\lambda_{i_1 - 1} < a < \lambda_{i_1} \le \lambda_{i_2} < b < \lambda_{i_2 + 1}$$

for certain indices i_1 and i_2. We are looking for a spectral projection onto the part of the spectrum $\sigma' = \sigma(M) \cap (a, b)$.

The ideal projection would be based on the function

$$\mathcal{X}(\lambda) = \begin{cases} \lambda & \text{for } \lambda \in (a, b), \\ 0 & \text{otherwise.} \end{cases} \tag{13.7}$$

An approximation of $\mathcal{X}$ is given by the rational function $\tilde{\mathcal{X}} : \mathbb{R} \to \mathbb{R}$ defined by

$$\tilde{\mathcal{X}}(\lambda) = \frac{\lambda}{1 + T(\lambda)^{2^\ell}} \qquad \text{with} \quad T(\lambda) = \frac{2\lambda - (b+a)}{b-a}$$

for $\ell \in \mathbb{N}$. For large ℓ, $\tilde{\mathcal{X}}$ is a good approximation of $\mathcal{X}$. To see this, we observe that for $a < \lambda < b$, $|T(\lambda)| < 1$. Thus, for large ℓ, $\tilde{\mathcal{X}}(\lambda) \approx \lambda$ in (a, b). On the other hand, if $\lambda > b$ or $\lambda < a$, $|T(\lambda)| > 1$ holds, yielding $\tilde{\mathcal{X}}(\lambda) \approx 0$ for large ℓ.

The function $\tilde{\mathcal{X}}$ can be extended to matrices in a straightforward way (cf. §14):

$$T(\mathbf{A}) = \frac{1}{b-a}(2\mathbf{A} - (b+a)\mathbf{I}), \quad \tilde{\mathcal{X}}(\mathbf{A}) = (\mathbf{I} + T(\mathbf{A})^{2^\ell})^{-1}\mathbf{A}\,.$$

The special structure of $\tilde{\mathcal{X}}$ is chosen, because $\tilde{\mathcal{X}}(\mathbf{A})$ can be computed with only ℓ matrix-matrix multiplications and one matrix inversion (another polynomial of high degree but few matrix-matrix multiplications has appeared in (14.11)). Again, if we apply methods requiring only matrix-vector multiplications, an LU decomposition of $\mathbf{I} + T(\mathbf{A})^{2^\ell}$ is sufficient. For details and numerical results, we refer to [148].

13.9 $\mathcal{H}$-AMLS

A very promising method is based on the 'component mode synthesis' (CMS; cf. Hurty [162]), whose recursive version is called AMLS (automated multi-level substructuring; cf. Bennighof [38]). It becomes an efficient method as soon as the hierarchical matrix arithmetic is applied. The resulting method is described in detail by Gerds–Grasedyck [95].

Let A be the finite element matrix discretising an elliptic boundary value problem defined by the bilinear form $a(\cdot,\cdot)$ with the underlying domain $\Omega \subset \mathbb{R}^d$ and the finite element space V_h. Decompose Ω as in §12.1 into disjoint subdomains Ω_i ($i = 1, 2$) with the interior boundary $\gamma = \partial\Omega_1 \cap \partial\Omega_2$ consisting of faces of the finite elements. Then V_h is the direct sum of the subspaces V_i corresponding to Ω_i and the subspace $V_\gamma = \{v \in V_h : a(v, w) = 0 \text{ for all } w \in V_h \text{ with } w|_\gamma = 0\}$. V_γ is the space of the discrete a-harmonic functions and arbitrary values on γ.

Let $n = \dim(V_h)$. We want to compute the first n_{ev} eigenfunctions of A where $n_{ev} \leq \mathcal{O}(\sqrt[3]{n})$. These eigenfunctions span the space E_Ω. Define E_{Ω_i} similarly for the subdomains Ω_i, while E_γ is spanned by the eigenfunctions of $a(\cdot,\cdot) : V_\gamma \times V_\gamma \to \mathbb{R}$. It turns out that E_Ω can be well approximated by $E := E_{\Omega_1} + E_{\Omega_2} + E_\gamma$. E is a $3n_{ev}$-dimensional space. Standard algorithms for solving the complete eigenvalue problem have cubic cost, but $(3n_{ev})^3 \leq \mathcal{O}(n)$ is still linear in n. The concrete computation requires the spanning (local) eigenfunctions to be known. Computing the eigenfunctions in Ω_i ($i = 1, 2$) is obtained by the recursive application of the described method. As soon as the repeated domain decomposition produces a subdomain with few degrees of freedom, a direct eigenvalue solver can be applied.