The Method of Conjugate Gradients for Solving Lyapunov Matrix Equation

Jana Lungová and Martin Plešinger

Department of Mathematics, Technical University of Liberec

{jana.lungova, martin.plesinger}@tul.cz

Motivation: The controll theory

In controll theory, we have physical systems represented by differential equations. We will focus on linear state space model represented by equations

$$-\frac{dx}{dt} = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t),$$

where x is a state vector, u is an input vector, y is a (measured) output vector and A, B, C and D are system matrices.

The goal is usually to get some specific output (or get the system to be in some specific state, which we asses by the measured output). To reach this, we need the system to be *controllable* and *observable*.

For stable systems the controllability and observability are determined using controllability and observability grammians

$$W_c = \int_0^\infty e^{-At}BB^Te^{-A^Tt}dt, \quad W_o = \int_0^\infty e^{-At}C^TCe^{-A^Tt}dt.$$

The grammians satisfy so-called Lyapunov equations

$$AW_c + W_cA^T = BB^T, \quad AW_o + W_oA^T = C^TC.$$

The Lyapunov equation

The Lyapunov equation we are focused on is

$$AX + XA^T = F.$$

If the marices A and F are symmetrical, than the sollution X also has to be symmetrical. We want to use the CG method, so we choose matrix A symmetrical and positive definite (SPD). If the right-hand side is of low rank, then the sollution can be approximated by a low rank matrix. This is important, because we want to simulate use of low-rank arithmetics.

The numerical rank of calculated sollutions X_k might increase to the point, where the computation would collapse due to unfullfilable memory demands.

Basic properties of matrices in the Lyapunov equation

It is usefull to rewrite the Lyapunov equation into the usual form of a system of linear algebraic equations. We introduce a binary matrix operation, the so-called Kronecker product, and a vectorization operation. We use those to rewrite the equation $AX + XA^T = F$ as

$$Lx = f, \quad \text{where} \quad L = A \otimes I + I \otimes A, \quad f = \mathrm{vec}(F).$$

Since our matrix A is SPD, the matrix $L = AX + XA^T$ is also symmetric positive definite, see [3], so the CG method can be used. The matrix F is of low rank, specifically of rank one, so it can be written as the product of two vectors $F = bb^T$. Using the newly introduced operations, we can write $\mathrm{vec}(F) = b \otimes b$.

The matrix formulation of the conjugate gradient method (MCG)

The algorithm is given SPD matrix, right-hand side, initial estimate (std. $X_0 = 0$) and runs as follows.

MCG for solving matrix Lyapunov equation $AX + XA^T = F$

1. **input** A, F, X_0
2. $R_0 \leftarrow F - AX_0 - X_0A^T$ — initial residual
3. $\rho_0 \leftarrow \langle R_0, R_0 \rangle$
4. $P_0 \leftarrow R_0$ — initial residual vector
5. **for** $k = 1, 2, \ldots$
6. $\quad W_k \leftarrow AP_{k-1} + P_{k-1}A^T$
7. $\quad \alpha_k \leftarrow \rho_{k-1} / \langle W_k, P_{k-1} \rangle$
8. $\quad \beta_k \leftarrow \rho_{k-1}$
9. $\quad X_k \leftarrow X_{k-1} + P_{k-1}\alpha_k$ — sollution approximation
10. $\quad R_k \leftarrow R_{k-1} - W_k\alpha_k$ — calculated residual
11. $\quad \rho_k \leftarrow \langle R_k, R_k \rangle$
12. $\quad \beta_k \leftarrow \rho_k / \beta_k$
13. $\quad P_k \leftarrow R_k + P_{k-1}\beta_k$ — direction vector
14. **endfor**
15. **return** X_k

The basic numerical experiment

For numerical experiments we used the *minus 1D discretization of the Laplace operator with a dirichlet boundary condition*, so the matrix A is sparse, symmetrical and positive definite. For the basic experiment we used $n = 100$ and for the right-hand side we used matrix of ones, which satisfies the requirement for symmetry and low rank, so

$$F = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} = bb^T \in \mathbb{R}^{100\times100},$$

$$A = \begin{bmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{100\times100}.$$

The matrix L is under these assumptions *minus 2D discretization of the Laplace operator with Dirichlet boundary condition* and is $L = (A \otimes I + I \otimes A) =$

$$\left[\begin{array}{cc|c|cc} 4 & -1 & & -1 & \\ -1 & \ddots & \ddots & & \ddots \\ & \ddots & \ddots & -1 & & \ddots \\ & & -1 & 4 & & & -1 \\ \hline & & & \ddots & & & \\ -1 & & & & & & \\ & \ddots & & & & & \\ & & \ddots & & \ddots & & \\ & & & -1 & & & \\ \hline & & & & -1 & & \\ & & & & & \ddots & \\ & & & & \ddots & & \ddots \\ & & & & & \ddots & \ddots & -1 \\ \hline & & & -1 & & & 4 & -1 \\ & & & & \ddots & & -1 & \ddots & \ddots \\ & & & & & \ddots & & \ddots & \ddots & -1 \\ & & & & & & -1 & & -1 & 4 \end{array}\right] \in \mathbb{R}^{100^2\times100^2}.$$

The matrix L has the same important properties as the matrix A (L is also symmetrical and positive deinite), so we can use the CG method for solving equation $Lx = f$.

The rank development of the calculated solution

Here, we measure the numerical rank by the number of eigenvalues that are in absolute value greater than the machine precision multiplied by the (in absolute value) largest eigenvalue of the matrix and n, i.e. greater than $\epsilon_M|\lambda_1|n$.

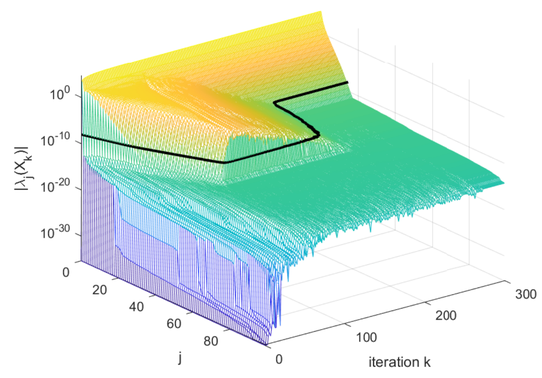

Figure 1: The graph shows the absolute values of (all 100 of) the eigenvalues of the matrices X_k (where k is an iteration). The black curve shows the progression of the measured rank of the matrix X_k.

Is the rank progress just a coincidence, or is the rank allways restricted?

We also see the same progression of rank in other conducted experiments, so it seems not to be coincidental. The goal now is to explain this phenomenon. The explenation should somehow be found in the matrices and the method used for solving the equation.

How many steps of CG do we need to execute?

The upper bound on the number of steps of CG (when solving $Lx = f$) is the number of distinct eigenvalues of the matrix L.

Eigenvalues of the matrix L

The properties of matrix L follow from the properties of matrix A. Let's say that the matrix A has k different eigenvalues ($|\mathrm{sp}(A)| = k$), let's denote them λ_i. Eigenvalues of the matrix L are all of the form $\lambda_i + \lambda_j$ ($i, j \in 1, \ldots k$, the case $i = j$ is also valid). Let's denote κ as the number of distinct eigenvalues of matrix L, κ is equal to the number of different sums we get, which is at most $\frac{k(k+1)}{2}$. Therefore $\kappa \le \frac{k(k+1)}{2}$.

Eigenvectors of the matrix L

The number of steps after which the conjugate gradient method (when solving equation $Dx = b$) converges to an exact solution (in exact arithmetic) is equal to the number of distinct eigenvalues of the matrix D such that the right-hand side b has a nonzero component in the corresponding eigenspace (there exists at least one eigenvector v corresponding to a given eigenvalue such that $\langle b, v \rangle \neq 0$).

We are solving $Lx = f$, where $f = b \otimes b$. The properties of matrix L follow from the properties of matrix A. We denoted r the eigenvectors of matrix A, so eigenvectors of matrix L are $r_i \otimes r_j$,

Let κ denote the number of distinct eigenvalues with the property that there exists an eigenvector to it that has a non-zero scalar product with the right side, where the spectrum $\mathrm{sp}(A)$ is the set of (distinct) eigenvalues of the matrix A. From the previous $|\mathrm{sp}(A)| = k$, from the condition on the corresponding eigenvector $\kappa \le k$.

If there is no eigenvector to the eigenvalue λ_d with a non-zero product with the right side, then for the matrix L and all its eigenvalues of the form $\lambda_d + \lambda_i$ the products $\langle b, r_d \rangle \langle b, r_i \rangle$ are necessarily zero. Similar to the number of eigenvalues of a matrix L, the number of distinct eigenvalues with corresponding eigenvectors $\langle b, r \rangle \neq 0$ is less than or equal to $\frac{\kappa(\kappa+1)}{2}$.

The eigenvectors in our numerical experiment

In the numerical exeriment we took minus discretization of 1D Laplace operator with Dirichlet boundary condition as the matrix $A \in \mathbb{R}^{n\times n}$. Eigenvectors r_j and eigenvalues λ_j of the matrix A are

$$r_j = \left[\sin\left(\frac{\rho j\pi}{n+1}\right)\right]_{\rho=1}^{n}, \quad \lambda_j = 2 - 2\cos\left(\frac{j\pi}{n+1}\right), \quad j = 1, \ldots, n,$$

for proof see e.g. [4]. The function cosine is monotone on interval $0 \le \frac{1}{n+1} \le x \le \frac{n}{n+1} < 1$, all eigenvalues are therefore distinct and we can focus only on the condition for eigenvectors.

On the right-hand side we had matrix of all ones $F = bb^T$, where b is vector of all ones ($b = [1, \ldots, 1]^T$). We need to find out the number of nonzero dot products $\langle b, r_j \rangle \neq 0$. The dot products are

$$\langle b, r_j \rangle = \sum_{\rho=1}^{n} \sin\left(\frac{\rho j\pi}{n+1}\right),$$

so we need to find out for how many eigenvectors their elements sum (subtract) to zero, the following image will help us with this.

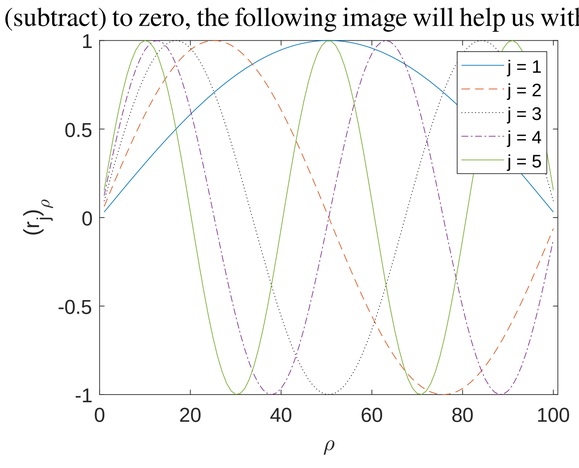

Figure 2: For clarity, the figure shows only the first five eigenvectors of the matrix $A \in \mathbb{R}^{n\times n}$, where $n = 100$.

For even and odd j we get different results, for odd indices the elements sum to some nonzero number, for all even indices the elements of the vectors sum to zero.

The right-hand side is therefore orthogonal to all eigenvectors r_j with even indices. Thus the (numerical) rank is restricted by the number of even numbers in the set $1, \ldots, n$, which is $\frac{n}{2}$ for even n and by $\frac{n-1}{2}$ for odd n. This explains the rank restriction observable in figure 1.

Theorem When solving Lyapunov equation using MCG method the following applies to the rank of matrix X_k (similarly for the rest of the matrices P_k, R_k, and W_k) in each step. Rank of the matrix X_k is bounded from above by the number of non-zero components b in the eigenspaces of matrix A,

$$\mathrm{rank}(X_k) \le \frac{\kappa(\kappa+1)}{2}.$$

References

[1] Daniel Kressner, Martin Plešinger, Christine Tobler: *A preconditioned low-rank CG method for parameter-dependent Lyapunov matrix equations*, Numerical Linear Algebra with Applications, Volume 21 (2014), Issue 5, pp. 666–684.

[2] Jörg Liesen, Zdeněk Strakoš: *Krylov Subspace Methods*, Numerical Mathematics and Scientific Computations, Oxford University Press, Oxford, 2012–3.

[3] Jana Lungová: *Convergence of matrix-oriented Krylov-subspace methods in terms of rank*, Bachelor thesis, TU of Liberec, Liberec, 2023.

[4] Jana Lungová: *Selected properties of the conjugate gradient method when applied on a matrix Lyapunov equation*, Diploma thesis, TU of Liberec, Liberec, 2025.

[5] Thilo Penzl: *Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case*, Systems & Control Letters, Volume 40 (2000), Issue 2, pp. 139–144.

[6] Jennifer Scott, Miroslav Tůma: *Algorithms for Sparse Linear Systems*, Neas Center Series (NECES) Birkhäuser, Cham, 2023.

[7] Valeria Simoncini, Yue Hao: *Analysis of the truncated conjugate gradient method for linear matrix equations*, SIAM Journal on Matrix Analysis and Applications, Volume 44 (2023), Issue 1, pp. 359–381.

[8] Zdeněk Strakoš, Petr Tichý: *On error estimation in the conjugate gradient method and why it works in finite precision computations*, Electronic Transactions on Numerical Analysis, Volume 13 (2002), pp. 56–80.

Acknowledgements. This work was supported by internal grant SGS-2026-4602 of Technical University of Liberec.