$\boxed{1}$ Assume that $A \in \mathbb{R}^{m\times n}$ with $m > n$ has full rank and that $b \in \mathbb{R}^m$. Consider the following iteration (the *CGNR-method*):

```
Set $x \leftarrow 0$, $r \leftarrow b$, $z \leftarrow A^T r$, $p \leftarrow z$, $s \leftarrow \|z\|_2^2$;
while not yet converged do
    $w \leftarrow Ap\,;$
    $\alpha \leftarrow s/\|w\|_2^2\,;$
    $x \leftarrow x + \alpha p\,;$
    $r \leftarrow r - \alpha w\,;$
    $z \leftarrow A^T r\,;$
    $s_{\text{old}} \leftarrow s\,;$
    $s \leftarrow \|z\|_2^2\,;$
    $\beta \leftarrow s/s_{\text{old}}\,;$
    $p \leftarrow z + \beta p\,;$
end
```

a) Show that this algorithm converges to a solution of the least squares problem

$$\min_x \|Ax - b\|_2.$$

b) Denote by x_k the result after k iterations of this algorithm. Show that x_k minimises $\|Ax - b\|_2$ among all vectors $x \in \mathcal{K}_k(A^T A, A^T b)$.

Hint: Show that this algorithm implements the CG-method for the normal equations.

Possible solution:

a) We recall that the solution of the least squares problem $\min_x \|Ax - b\|_2$ can be found by solving the normal equations $A^T Ax = A^T b$. A direct application of the CG method for this equation (with $x_0 = 0$) would read as (we use the variable z for storing the residual):

```
Set $x \leftarrow 0$, $z \leftarrow A^T b$, $p \leftarrow z$, $s \leftarrow \|z\|_2^2$;
while not yet converged do
    $v \leftarrow A^T A p$;
    $\alpha \leftarrow s/(p,v)$;
    $x \leftarrow x + \alpha p$;
    $z \leftarrow z - \alpha v$;
    $s_{\text{old}} \leftarrow s$;
    $s \leftarrow \|z\|_2^2$;
    $\beta \leftarrow s/s_{\text{old}}$;
    $p \leftarrow z + \beta p$;
end
```

In the initialisation, the variables in this algorithm have the same values as in the algorithm proposed in the exercise. Moreover, we have

$$\|w\|_2^2 = (Ap, Ap) = (p, A^T Ap) = (p, v),$$

and thus the step length α in the two algorithms is computed in the same manner. Additionally, the variable z is updated in the CGNR method as

$$z \leftarrow A^T(r - \alpha w) = A^T r - \alpha A^T w = z - \alpha A^T Ap.$$

Since $A^T Ap = v$, also the update of z is computed in the same manner in both of the algorithms. Since all the other lines of the algorithms are identically, they produce the same results. Thus the proposed algorithm is simply a reformulation of the CG method for the normal equations.

b) As shown above, the proposed algorithm is an implementation of the CG method for the normal equations $A^T Ax = A^T b$. As a consequence, the iterates x_k solve the optimisation problems

$$\min_{x \in \mathcal{K}_k(A^T A, A^T b)} \|x - x^*\|_{A^T A}^2,$$

where x^* is the solution of $A^T Ax = A^T b$. Now note that

$$\|x - x^*\|_{A^T A}^2 = (A^T A(x - x^*), x - x^*) = (A(x - x^*), A(x - x^*)) = \|A(x - x^*)\|_2^2.$$

That is, the iterates x_k solve the problems

$$\min_{x \in \mathcal{K}_k(A^T A, A^T b)} \|A(x - x^*)\|_2^2.$$

Now note that x^* solves the problem

$$\min_x \|Ax - b\|_2^2,$$

which implies that Ax^* is the orthogonal projection of b onto the range of A. This, however, implies that $Ax^* - b$ is perpendicular to all vectors of the form $A(x - x^*)$. Thus

$$\begin{aligned}
\|Ax - b\|_2^2 &= \|A(x - x^*) + Ax^* - b\|_2^2 \\
&= \|A(x - x^*)\|_2^2 + 2(A(x - x^*), Ax^* - b) + \|Ax^* - b\|_2^2 \\
&= \|A(x - x^*)\|_2^2 + \|Ax^* - b\|_2^2
\end{aligned}$$

for all $x \in \mathbb{R}^n$, and therefore

$$\|A(x - x^*)\|_2^2 = \|Ax - b\|_2^2 - \|Ax^* - b\|_2^2$$

for all $x \in \mathbb{R}^n$. Since $\|Ax^* - b\|_2^2$ is a constant independent of x, this implies that minimising $\|A(x - x^*)\|_2^2$ over some subspace of $\mathbb{R}^n$ is equivalent to minimising $\|Ax - b\|_2^2$ over the same subspace. In particular, this shows that the iterates x_k actually solve the optimisation problems

$$\min_{x \in \mathcal{K}_k(A^T A, A^T b)} \|Ax - b\|_2^2.$$

$\boxed{2}$ (Cf. Exercise 9.10 in YS) We consider the solution of a linear system $Ax = b$ using the pre-conditioned GMRES method with some pre-conditioner M. In the lecture we have discussed left- and right-pre-conditioning for this method, and we have also briefly discussed a general convergence result for the GMRES method.

a) Show that the matrices AM^{-1} and $M^{-1}A$ have the same eigenvalues. How are the eigenvectors of the two matrices related to each other?

b) Using the results of part **a)**, would you expect that the left- and the right-pre-conditioned iterations converge:

1. …in exactly the same number of steps?
2. …in roughly the same number of steps?
3. …in roughly the same number of steps provided that the system is not ill-conditioned?

Possible solution

a) The characteristic polynomial of the left preconditioned matrix $M^{-1}A$ is the polynomial

$$p_\ell(\lambda) = \det(\lambda \operatorname{Id} - M^{-1}A) = \det(\lambda M^{-1}M - M^{-1}A) = \det(\lambda M^{-1}(M - A)).$$

Because of the multiplicativity of the determinant, we can switch the order of the two matrices, and thus obtain that

$$p_\ell(\lambda) = \det(\lambda(M - A)M^{-1}) = \det(\lambda \operatorname{Id} - AM^{-1}).$$

This, however, is precisely the characeristic polynomial of the right preconditioned matrix AM^{-1}. In particular, this shows that AM^{-1} and $M^{-1}A$ have the same eigenvalues.

Now assume that v is an eigenvector of AM^{-1}, that is,

$$AM^{-1}v = \lambda v.$$

Then

$$M^{-1}AM^{-1}v = \lambda M^{-1}v,$$

which shows that $M^{-1}v$ is an eigenvector of $M^{-1}A$.

(Alternatively, you can also verify that $AM^{-1} = MM^{-1}AM^{-1} = M(M^{-1}A)M^{-1}$ is a similarity transform of $M^{-1}A$.)

b) Unless the matrices A and M^{-1} commute, in which case the right and the left preconditioned systems are equivalent, the iterates for the right and the left preconditioned systems will be different, and therefore we cannot expect that they converge in precisely the same number of steps. The best we can hope for is that the number of steps is almost the same. In order to see whether this will be the case, we use the error estimate from Saad, Proposition 6.32:

Assuming that the matrix AM^{-1} is diagonalisable, that is,

$$AM^{-1} = X\Lambda X^{-1},$$

we have for the right preconditioned system the estimate

$$\frac{\|r_m^{(r)}\|_2}{\|r_0\|_2} \le \kappa_2(X)\epsilon_m,$$

where

$$\epsilon_m = \min_{p\in\mathbb{P}_m,\, p(0)=1} \max_i |p(\lambda_i)|.$$

Now consider the left preconditioned system with system matrix $M^{-1}A$. If $AM^{-1} = X\Lambda X^{-1}$, then

$$M^{-1}A = M^{-1}X\Lambda X^{-1}M.$$

Thus we obtain for the left preconditioned system the estimate

$$\frac{\|r_m^{(r)}\|_2}{\|r_0\|_2} \le \kappa_2(M^{-1}X)\epsilon_m.$$

Now note that we can estimate

$$\kappa_2(M^{-1}X) \le \kappa_2(M^{-1})\kappa_2(X)$$

and

$$\kappa_2(X) \le \kappa_2(M)\kappa_2(M^{-1}X) = \kappa_2(M^{-1})\kappa_2(M^{-1}X).$$

That is

$$\frac{\kappa_2(M^{-1}X)}{\kappa_2(X)} \le \kappa_2(M^{-1}) \qquad \text{and} \qquad \frac{\kappa_2(X)}{\kappa_2(M^{-1}X)} \le \kappa_2(M^{-1}).$$

In particular, if the matrix M^{-1} is well-conditioned, that is $\kappa_2(M^{-1}) \approx 1$, the error estimates for the left and the right preconditioned systems are nearly the same, and we would expect the two different iterations to converge in roughly the same number of steps. If, however, the matrix M^{-1} is ill-conditioned, this is not the case, and the expected number of iterations based on the error estimates would be noticeably different.

Finally, we note that we have for a reasonable preconditioner that $M \approx A$. In particular, the matrix M can be expected to be ill-conditioned if (and only if) the matrix A is. In total, we can say that we would expect roughly the same number of iterations for both variants of the preconditioned system, as long as the system is well-conditioned.

$\boxed{3}$ We consider once again the solution of the one-dimensional Poisson problem

$$\begin{aligned} -u''(x) &= f(x), \quad 0 < x < 1, \\ u(0) = u(1) &= 0, \end{aligned}$$

discretised using finite differences on a uniform grid with step-size $h = 1/n$. Denote (again) the discretised equations as a system $Au = b$, where A is the discretised second derivative.

a) Suppose that we apply the CG method in order to solve this problem. Estimate how many steps will be needed in order to reduce the initial error (measured in the A-norm) by 5 orders of magnitude.

b) We now solve the same problem using the pre-conditioned CG method, and we choose as pre-conditioner two iterates of the Jacobi method. Verify that this pre-conditioner is symmetric, and estimate how many steps of the pre-conditioned method will be necessary to reduce the initial error by 5 orders of magnitude.

c) Determine whether this type of pre-conditioning makes sense in this situation.

Hint: If two matrices A and B have the same eigenvectors v_i with eigenvalues λ_i and μ_i, respectively, then the matrix AB will have the eigenvalues $\lambda_i \mu_i$. (Also, we have $AB = BA$). Also, if p is a polynomial, then the eigenvalues of $p(A)$ are the values $p(\lambda_i)$.

Possible solution:

The matrix of the discretised system we are solving is a tridiagonal Toeplitz matrix with diagonal entries 2, and offdiagonal entries -1. According to the note on the eigenvalues of tridiagonal Toeplitz matrices, this matrix has the eigenvalues

$$\lambda_m = 2\Big(1 - \cos\Big(\frac{m\pi}{n}\Big)\Big), \qquad m = 1, \ldots, n-1.$$

a) We can use the error estimate

$$\frac{\|x_k - x^*\|_A}{\|x_0 - x^*\|_A} \le 2\bigg(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}\bigg)^k$$

for the CG method. In order to estimate the condition of the matrix, we note that

$$\lambda_1 = 2\Big(1 - \cos\Big(\frac{\pi}{n}\Big)\Big) \approx \frac{\pi^2}{n^2}$$

and

$$\lambda_n = 2\Big(1 + \cos\Big(\frac{\pi}{n}\Big)\Big) \approx 4.$$

These are also the smallest and largest eigenvalue of the matrix of the system. Thus the condition of the matrix is

$$\kappa = \frac{\lambda_n}{\lambda_1} \approx \frac{4\pi^2}{n^2}$$

and

$$\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \approx \frac{1-\frac{2\pi}{n}}{1+\frac{2\pi}{n}}.$$

As a consequence, we have the error estimate

$$\frac{\|x_k - x^*\|_A}{\|x_0 - x^*\|_A} \lesssim 2\Bigl(\frac{1-\frac{2\pi}{n}}{1+\frac{2\pi}{n}}\Bigr)^k.$$

If we want to achieve a reduction of the error by a factor of at least 10^{-5}, we have to choose k such that

$$2\Bigl(\frac{1-\frac{2\pi}{n}}{1+\frac{2\pi}{n}}\Bigr)^k \le 10^{-5},$$

or, equivalently,

$$k \ln\Bigl(\frac{1-\frac{2\pi}{n}}{1+\frac{2\pi}{n}}\Bigr) \le \ln(10^{-5}/2).$$

Now we estimate

$$\ln\Bigl(\frac{1-\frac{2\pi}{n}}{1+\frac{2\pi}{n}}\Bigr) \approx -\frac{4\pi}{n},$$

and obtain the condition

$$k \ge \frac{\ln(2\cdot 10^5)}{4\pi} n.$$

b) We now use two iterates of the Jacobi method as a preconditioner. That is, the product $w = M^{-1}z$ is evaluated by performing two steps of the Jacobi method for solving $Aw = z$ with initialisation $w_0 = 0$. In this particular case, the diagonal of the matrix A is simply the constant 2, and thus the Jacobi iteration reads

$$w_{k+1} = \frac{1}{2}(D-A)w_k + \frac{1}{2}z = \Bigl(\mathrm{Id} - \frac{1}{2}A\Bigr)w_k + \frac{1}{2}z.$$

With $w_0 = 0$ we have

$$w_1 = \frac{1}{2}z$$

and thus

$$w_2 = \Bigl(\mathrm{Id} - \frac{1}{2}A\Bigr)w_1 + \frac{1}{2}z = \Bigl(\mathrm{Id} - \frac{1}{2}A\Bigr)\frac{1}{2}z + \frac{1}{2}z = \Bigl(\mathrm{Id} - \frac{1}{4}A\Bigr)z.$$

Thus

$$M^{-1} = \mathrm{Id} - \frac{1}{4}A.$$

Obviously, this matrix is symmetric. Also, we have seen above that all eigenvalues of A are smaller than 4, and therefore M^{-1} is positive definite.

In order to obtain error estimates for the preconditioned system, we require an estimate of the condition of the matrix $M^{-1}A = A - \frac{1}{4}A^2$ and then follow the same approach as in part **a)**. The eigenvalues of this matrix are

$$\mu_m = \lambda_m - \frac{1}{4}\lambda_m^2 \qquad \text{where} \qquad \lambda_m = 2\Bigl(1 - \cos\Bigl(\frac{m\pi}{n}\Bigr)\Bigr)$$

is an eigenvalue of A. This simplifies to

$$\mu_m = \sin^2\Bigl(\frac{m\pi}{n}\Bigr).$$

It is easy to see that the largest eigenvalue is either equal to or close to 1, whereas the smallest eigenvalue is

$$\mu_1 = \sin^2\Big(\frac{\pi}{n}\Big) \approx \frac{\pi^2}{n^2}.$$

As a consequence, the condition of the preconditioned matrix is approximately

$$\kappa(M^{-1}A) \approx \frac{n}{\pi},$$

and

$$\frac{\sqrt{\kappa(M^{-1}A)}-1}{\sqrt{\kappa(M^{-1}A)}+1} \approx \frac{1-\frac{\pi}{n}}{1+\frac{\pi}{n}}.$$

Since

$$\ln\Big(\frac{1-\frac{\pi}{n}}{1+\frac{\pi}{n}}\Big) \approx -\frac{2\pi}{n},$$

we conclude that

$$k \geq \frac{\ln(2\cdot 10^5)}{2\pi}n$$

iterates of the preconditioned CG method are necessary in order to obtain an error reduction of 5 orders of magnitude.

c) Based on the results of **a)** and **b)**, we know that the preconditioned CG method requires roughly half the number of iterations as the standard CG method. However, this comes at the additional cost of having to evaluate the preconditioner once in each iteration.

Let us first analyse the number of operations needed in the standard CG method. Here we see that each iteration requires:

- one evaluation of the operator A, which requires $5n$ operations,
- the computation of two inner products, each of which requires $2n$ operations,
- three products of a scalar with a vector, each of which requires n operations,
- three sums of two vectors, each of which requires n operations.

In total, each step of the standard CG method requires $15n$ operations.

Now we compare this to the preconditioned CG method. In this particular case, where the preconditioner $M^{-1} = \mathrm{Id} - \frac{1}{4}A$ is tridiagonal, its evaluation is as expensive as one evaluation of the operator A, which requires roughly $5n$ operations. That is, compared to the standard CG method, this particular preconditioned version requires an additional $5n$ calculations. Put differently, the number of operations increases by a factor of $4/3$. Since only half the number of iterations are required, the total cost for obtaining the same accuracy as with the CG method is reduced by a factor of $2/3$. Therefore the preconditioned method makes sense, if we want to decrease computation time.

(Note, however, that this particular system can be solved in linear time with a standard LU or Cholesky factorisation.)