

Conjugate gradient method with preprocessing (CG method) *

Nobuyuki Umetani

March 23, 2021

Contents

1	CG method with preprocessing (PCG method)	1
1.1	algorithm (PCG method)	3
2	Preprocessing method	3

1 CG method with preprocessing (PCG method)

In the previous section, we found that the convergence of the CG method strongly depends on the eigenvalue distribution of the matrix \mathbf{A} . Let us suppose that the equation $\mathbf{Ax} = \mathbf{b}$ to be solved can be transformed as follows using the nonspecific matrix \mathbf{P} .

$$\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}} \quad (1)$$

However, $\tilde{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}$, $\tilde{\mathbf{x}} = \mathbf{P}^T\mathbf{x}$, $\tilde{\mathbf{b}} = \mathbf{P}^{-1}\mathbf{b}$.

When \mathbf{A} is positive definite symmetry, $\tilde{\mathbf{A}}$ is also positive definite symmetric. Consider applying CG method to simultaneous linear equation $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$. If $\tilde{\mathbf{A}}$ has a distribution of good eigenvalues such that the condition number is smaller than \mathbf{A} , we can see that the solution can be obtained more quickly. Here we define a matrix $\mathbf{M} = \mathbf{P}\mathbf{P}^T$. In the following, without using the matrix \mathbf{P} , $\tilde{\mathbf{A}}$ or $\tilde{\mathbf{b}}$, rewrite the CG algorithm so that it is the same as applying the CG method to the simultaneous linear equation $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$, using only the inverse matrix \mathbf{M}^{-1} of \mathbf{M} . When it is $\mathbf{M} = \mathbf{A}$, it becomes $\tilde{\mathbf{A}} = \mathbf{I}$ and a solution

*This is a memorandum to write down what a forgetful author studied a long time ago. Surely it contains many mistakes. Excuse me. It is appreciated if you let me know if you have any comments or suggestions.

is required without solving the deformed simultaneous linear equation $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$. Let $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{p}}$ be the residuals $\tilde{\mathbf{r}}$ and the search direction vector $\tilde{\mathbf{p}}$ when CG method is applied to the simultaneous linear equation $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$.

$$\tilde{\mathbf{r}} = \tilde{\mathbf{b}} - \tilde{\mathbf{A}}\tilde{\mathbf{x}} \quad (2)$$

$$= \mathbf{P}^{-1}\mathbf{b} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}\mathbf{P}^T\mathbf{x} \quad (3)$$

$$= \mathbf{P}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}) \quad (4)$$

$$= \mathbf{P}^{-1}\mathbf{r} \quad (5)$$

$$(\tilde{\mathbf{r}}, \tilde{\mathbf{r}}) = (\mathbf{P}^{-1}\mathbf{r}, \mathbf{P}^{-1}\mathbf{r}) = (\mathbf{r}, \mathbf{P}^{-T}\mathbf{P}^{-1}\mathbf{r}) = (\mathbf{r}, \mathbf{M}^{-1}\mathbf{r}) \quad (6)$$

$$(\tilde{\mathbf{p}}, \tilde{\mathbf{A}}\tilde{\mathbf{p}}) = (\tilde{\mathbf{p}}, \mathbf{P}^{-1}\mathbf{A}\mathbf{P}^{-T}\tilde{\mathbf{p}}) = (\mathbf{P}^{-T}\tilde{\mathbf{p}}, \mathbf{A}\mathbf{P}^{-T}\tilde{\mathbf{p}}) \quad (7)$$

Therefore, if $\mathbf{z} = \mathbf{M}^{-1}\mathbf{r}$, $\mathbf{p} = \mathbf{P}^{-T}\tilde{\mathbf{p}}$ are given, the coefficients of the CG method can be expressed as follows.

$$\alpha_k = \frac{(\tilde{\mathbf{r}}_k, \tilde{\mathbf{r}}_k)}{(\tilde{\mathbf{p}}_k, \tilde{\mathbf{A}}\tilde{\mathbf{p}}_k)} \quad (8)$$

$$= \frac{(\mathbf{r}_k, \mathbf{z}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} \quad (9)$$

$$\beta_k = \frac{(\tilde{\mathbf{r}}_{k+1}, \tilde{\mathbf{r}}_{k+1})}{(\tilde{\mathbf{r}}_k, \tilde{\mathbf{r}}_k)} = \frac{(\mathbf{r}_{k+1}, \mathbf{z}_{k+1})}{(\mathbf{r}_k, \mathbf{z}_k)} \quad (10)$$

From the relational expression of the CG method,

$$\Delta\tilde{\mathbf{x}}_k = \tilde{\mathbf{x}}_{k+1} - \tilde{\mathbf{x}}_k = \alpha_k\tilde{\mathbf{p}}_k = \alpha_k\mathbf{P}^T\mathbf{p}_k \quad (11)$$

$$\Delta\tilde{\mathbf{r}}_k = \tilde{\mathbf{r}}_{k+1} - \tilde{\mathbf{r}}_k = -\tilde{\mathbf{A}}\Delta\tilde{\mathbf{x}}_k = -\alpha_k\tilde{\mathbf{A}}\tilde{\mathbf{p}}_k = -\alpha_k\mathbf{P}^{-1}\mathbf{A}\mathbf{p}_k \quad (12)$$

$$\tilde{\mathbf{p}}_{k+1} = \tilde{\mathbf{r}}_{k+1} + \beta_k\tilde{\mathbf{p}}_k \quad (13)$$

Is satisfied. Therefore, from now on, we can obtain the solution \mathbf{x} , the residual \mathbf{r} , the update of the vector \mathbf{p} .

$$\Delta\mathbf{x}_k = \mathbf{P}^{-T}\Delta\tilde{\mathbf{x}}_k = \alpha_k\mathbf{P}^{-T}\mathbf{P}^T\mathbf{p}_k = \alpha_k\mathbf{p}_k \quad (14)$$

$$\Delta\mathbf{r}_k = \mathbf{P}\Delta\tilde{\mathbf{r}}_k = -\alpha_k\mathbf{P}\mathbf{P}^{-1}\mathbf{A}\mathbf{p}_k = -\alpha_k\mathbf{A}\mathbf{p}_k \quad (15)$$

$$\mathbf{p}_{k+1} = \mathbf{P}^{-T} \tilde{\mathbf{p}}_{k+1} = \mathbf{P}^{-T} \tilde{\mathbf{r}}_{k+1} + \beta_k \mathbf{P}^{-T} \tilde{\mathbf{p}}_k \quad (16)$$

$$= \mathbf{P}^{-T} \mathbf{P}^{-1} \mathbf{r}_{k+1} + \beta_k \mathbf{P}^{-T} \mathbf{P}^T \mathbf{p}_k \quad (17)$$

$$= \mathbf{M}^{-1} \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k = \mathbf{z}_{k+1} + \beta_k \mathbf{p}_k \quad (18)$$

Therefore, the algorithm of the CG method is as follows.

1.1 algorithm (PCG method)

1. Compute $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$, $\mathbf{z}_0 = \mathbf{M}^{-1}\mathbf{r}_0$, $\mathbf{p}_0 = \mathbf{z}_0$

2. For $k = 0, 1, \dots, m$ Do :

$$(a) \alpha_k = \frac{(\mathbf{r}_k, \mathbf{z}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)}$$

$$(b) \mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$$

$$(c) \mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A}\mathbf{p}_k$$

$$(d) \mathbf{z}_{k+1} = \mathbf{M}^{-1} \mathbf{r}_{k+1}$$

$$(e) \beta_k = \frac{(\mathbf{r}_{k+1}, \mathbf{z}_{k+1})}{(\mathbf{r}_k, \mathbf{z}_k)}$$

$$(f) \mathbf{p}_{k+1} = \mathbf{z}_{k+1} + \beta_k \mathbf{p}_k$$

3. End Do

2 Preprocessing method

\mathbf{M} is called a preprocessing matrix. It is not the preprocessing matrix \mathbf{M} itself that is actually used, but the preprocessing matrix \mathbf{M} can only solve the simultaneous linear equation $\mathbf{M}\mathbf{z} = \mathbf{r}$ to find \mathbf{z} . If solving \mathbf{M} does not change much difficulty than solving \mathbf{A} , it is totally overwhelmed. That is, the preprocessing matrix \mathbf{M} needs to be able to solve more easily than \mathbf{A} . The preprocessing matrix \mathbf{M} is required to have the property $\mathbf{M} \simeq \mathbf{A}$ which is sufficiently close to \mathbf{A} . When it is $\mathbf{M} = \mathbf{A}$, it becomes $\tilde{\mathbf{A}} = \mathbf{I}$ and it converges on only one iteration. In other words, if $\mathbf{M} \simeq \mathbf{A}$, $\tilde{\mathbf{A}} \simeq \mathbf{I}$ would be a good eigenvalue distribution and convergence would be fast. Also, in terms of the CG method, since it was expressed as $\mathbf{M} = \mathbf{P}\mathbf{P}^T$, the pre-processing matrix \mathbf{M} must be positive definite symmetric. Conversely, when \mathbf{M} is positive definite symmetric, it is known that there is a real matrix \mathbf{P} that becomes $\mathbf{M} = \mathbf{P}\mathbf{P}^T$.

In summary,

desirable property as pre-processing matrix

The preprocessing matrix \mathbf{M} solves $\mathbf{Mz} = \mathbf{r}$ and can easily find \mathbf{z} , a positive definite symmetric matrix approximating \mathbf{A}

It can be said that. A representative example of preprocessing in which the preprocessing matrix for the CG method is symmetric is

- Diagonal Scaling
- Incomplete LU factorization (ILU factorization, Incomplete LU Factorization)
- SSOR method (Symmetric Successive Over Relaxation)
- Point Jacobi Method (Point Jacobi Method)
- Multigrid Method (Multigrid Method)

The PCG method using incomplete LU decomposition as preprocessing is often called ICCG method in particular and is often used.