Conjugate Gradient Method

Nobuyuki Umetani

July 14, 2018

Contents

1 Introduction

The CG method (Conjugate Gradient Methods) is a method proposed by M. R. Hestenes and E. Stiefel in 1952 cite hestenes - methods

The CG method is a method for solving linear equations used for positive definite symmetric matrices by an iterative method.

1.0.1 Positive Definite Matrix

We write the inner product of vector u, v like (u, v). The real matrix $\mathbf{A}$ is positive definite if

$$(\mathbf{Au}, \mathbf{u}) \geq 0 \quad \forall \mathbf{u} \in \mathbf{R}^n \quad (\ equality\ holds\ only\ if\ \mathbf{u} = 0 \quad). \tag{1}$$

It means that $\mathbf{A}$ is symmetric

$$(\mathbf{Aa}, \mathbf{b}) = (\mathbf{a}, \mathbf{Ab}) \quad \forall \mathbf{a}, \mathbf{b} \in \mathbf{R}^n. \tag{2}$$

2 Basics of CG method

Now, let's solve the linear homogeneous equation $\mathbf{Ax} = \mathbf{b}$ as follows.

In the k th iteration of the CG method, the $\mathbf{A}$ norm of the error defined using the solution $\mathbf{x}$ and error $\mathbf{e}$ of this equation as follows

$$||\mathbf{e}||_A^2 = (\mathbf{e}, \mathbf{Ae}) = ||\mathbf{x}_k - \mathbf{x}||_A^2 = (\mathbf{x}_k - \mathbf{x}, \mathbf{A}(\mathbf{x}_k - \mathbf{x})) \geq 0 \quad (equality\ holds\ only\ if\ \mathbf{x}_k = \mathbf{x}) \tag{3}$$

This is a method to find an approximate solution $\mathbf{x}_k$ that minimizes the subspace $\mathcal{K}_k + \mathbf{x}_0$. However, $\mathcal{K}_k$ is the Krylov subspace $\mathcal{K}_k = span\{\mathbf{r}_0, \mathbf{Ar}_0, \mathbf{A}^2\mathbf{r}_0, \cdots, \mathbf{A}^{k-1}\mathbf{r}_0\}$.

That is, the CG method is a method for finding an approximate solution $\mathbf{x}_k$ of simultaneous linear equations as follows.

begin screen

$$find \quad \mathbf{x}_k \in \mathbf{x}_0 + \mathcal{K}_k \quad that \quad minimize||\mathbf{x} - \mathbf{x}_k||_A \tag{4}$$

end screen

In this way, when the distance represented by the $\mathbf{A}$ norm with $\mathbf{x}$ in the subspace $\mathcal{K}_k + \mathbf{x}_0$ takes an extreme value, the following orthogonality relation clearly holds.

$$(\mathbf{x}_k - \mathbf{x}, \mathbf{w})_\mathbf{A} = 0 \qquad \forall \mathbf{w} \in \mathcal{K}_k \tag{5}$$

However, it is defined as the inner product $(\mathbf{a}, \mathbf{b})_\mathbf{A} = (\mathbf{Aa}, \mathbf{b})$ defined by the matrix. Using this, the CG method can be expressed as follows.

begin screen

$$find \quad \mathbf{x}_k \in \mathbf{x}_0 + \mathcal{K}_k \qquad so \quad that \qquad \mathbf{x}_k - \mathbf{x} \quad \perp_\mathbf{A} \quad \mathcal{K}_k \tag{6}$$

end screen

However, it shows the following relationship with $\perp_\mathbf{A}$.

$$\mathbf{a} \quad \perp_\mathbf{A} \quad \mathbf{b} \Leftrightarrow (\mathbf{a}, \mathbf{b})_\mathbf{A} = 0 \Leftrightarrow (\mathbf{Aa}, \mathbf{b}) = 0 \Leftrightarrow \mathbf{Aa} \perp \mathbf{b} \tag{7}$$

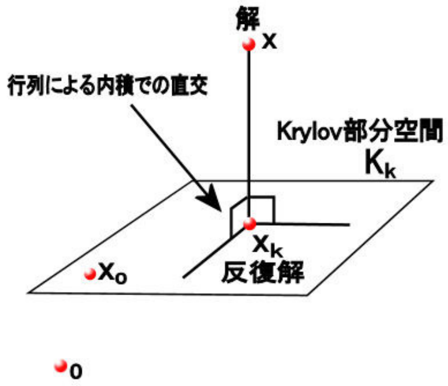

Figure 1: The iteration of conjugate gradient method(In Eucledian space)

In other words, the CG method can be said to be a method of orthogonally projecting the solution $\mathbf{x}$ to the subspace $\mathcal{K}_k + \mathbf{x}_0$ in the inner product defined by the matrix. Krylov subspace $\mathcal{K}_k$ is a finite dimensional subspace, so it is a complete linear space. Thus, from the Lax-Milgram theorem, such projection always exists.

Also, since the residual $\mathbf{r}_k$ at k times iteration is $\mathbf{r}_k = \mathbf{A}(\mathbf{x}_k - \mathbf{x})$, the CG method can be said to find a solution as follows.

$$find \quad \mathbf{x}_k \in \mathbf{x}_0 + \mathcal{K}_k \qquad so \quad that \qquad \mathbf{r}_k \quad \perp \quad \mathcal{K}_k \tag{8}$$

This means that the solution is looked for in the $\mathcal{K}_k + \mathbf{x}_0$ so that the residual $\mathbf{r}_k$ is orthogonal to the subspace $\mathcal{K}_k$.

The CG method has a deep relationship with the Lanczos method, which is a method of creating an orthonormal basis of Krylov subspace.

3 Basic procedure

3.1 increment of solution

Suppose that the increment of the solution from the k iterative solution $\mathbf{x}_k$ to the $k+1$ iterative solution $\mathbf{x}_{k+1}$ is written using the coefficient α_k and the vector $\mathbf{p}_k$ as follows.

$$\mathbf{x}_{k+1} - \mathbf{x}_k = \alpha_k \mathbf{p}_k \tag{9}$$

Since $\mathbf{p}$ determines the direction of solution increment, it is called search direction vector.

Updating the solution to the $k+2$ iterative solution $\mathbf{x}_{k+2}$ The vector $\mathbf{p}_{k+1}$ is determined as follows using the residual $\mathbf{r}_{k+1}$ of the solution $\mathbf{x}_{k+1}$, the coefficient β_k, and the update vector $\mathbf{p}_k$ of the previous solution as follows

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k \tag{10}$$

How to determine the coefficients β, α is explained below.

3.2 Determination of coefficient β

From the previous discussion, in the CG method, the difference between the exact solution $\mathbf{x}$ and the iterative solution $\mathbf{x}_{k+1}$ was orthogonal in the $\mathbf{A}$ norm to the Krylov subspace $\mathcal{K}_{k+1}$. In other words,

$$\mathbf{x}_{k+1} - \mathbf{x} \quad \perp_{\mathbf{A}} \quad \mathcal{K}_{k+1} \tag{11}$$

Therefore, since the difference $\mathbf{x}_{k+1} - \mathbf{x}$ from the exact solution lies in the orthogonal space in the Krylov subspace $\mathcal{K}_{k+1}$ and $\mathbf{A}$ norm, update the solution update vector $\mathbf{p}_{k+1}$ to the next iterative solution $\mathbf{x}_{k+2}$ from the orthogonal space with $\mathbf{A}$ norm to this $\mathcal{K}_{k+1}$ $\mathbf{x}_{k+2}$ should be closer to the exact solution.

As shown later, the previous search direction vector $\mathbf{p}_k$ of $\mathbf{p}_{k+1}$ is in the subspace $\mathcal{K}_{k+1}$ trying to make it orthogonal. In other words,

$$\mathbf{p}_k \in \mathcal{K}_{k+1} \tag{12}$$

Therefore, in order for the search direction vector $\mathbf{p}_{k+1}$ to be orthogonal to $\mathcal{K}_{k+1}$ with the $\mathbf{A}$ norm, it is necessary to orthogonalize to the minimum $\mathbf{p}_k$ with $\mathbf{A}$ norm. In other words,

begin screen

$$(\mathbf{p}_{k+1}, \mathbf{p}_k)_{\mathbf{A}} = (\mathbf{p}_{k+1}, \mathbf{A}\mathbf{p}_k) = 0 \tag{13}$$

end screen

As shown later, the search direction chosen as above is orthogonal not only to $\mathbf{p}_k$ but also to the $\mathcal{K}_{k+1}$ space by the $\mathbf{A}$ norm.

An expression that defines a search direction vector from the above residual and the previous search direction vector,

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k \tag{14}$$

, It means that a new search direction vector $\mathbf{p}_{k+1}$ can be obtained by projecting the residual difference $\mathbf{r}_{k+1}$ to the orthogonal space with the Krylov subspace $\mathcal{K}_{k+1}$ and $\mathbf{A}$ norm.

From this equation, we can determine β.

Substituting this expression into the expression $(\mathbf{p}_{k+1}, \mathbf{A}\mathbf{p}_k) = 0$ where the two search directions are orthogonal, the parameter β becomes as follows.

$$\beta = -\frac{(\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} \tag{15}$$

Furthermore, using the relational expression $(\mathbf{r}_i, \mathbf{r}_j) = 0 \quad (\quad i \neq j \quad)$ of orthogonality of residuals to be described later,

$$(\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_k) = \left(\mathbf{r}_{k+1}, -\frac{1}{\alpha_k}(\mathbf{r}_{k+1} - \mathbf{r}_k)\right) = -\frac{1}{\alpha_k}(\mathbf{r}_{k+1}, \mathbf{r}_{k+1}) \tag{16}$$

$$(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k) = (\mathbf{r}_k + \beta_{k-1}\mathbf{p}_{k-1}, \mathbf{A}\mathbf{p}_k) = (\mathbf{r}_k, \mathbf{A}\mathbf{p}_k) = \left(\mathbf{r}_k, -\frac{1}{\alpha_k}(\mathbf{r}_{k+1} - \mathbf{r}_k)\right) = \frac{1}{\alpha_k}(\mathbf{r}_k, \mathbf{r}_k) \tag{17}$$

As a result, the coefficient β_k can be further expressed as follows.

$$\beta_k = -\frac{(\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} = -\frac{-\frac{1}{\alpha_k}(\mathbf{r}_{k+1}, \mathbf{r}_{k+1})}{\frac{1}{\alpha_k}(\mathbf{r}_k, \mathbf{r}_k)} = \frac{(\mathbf{r}_{k+1}, \mathbf{r}_{k+1})}{(\mathbf{r}_k, \mathbf{r}_k)} \tag{18}$$

3.3 Determination of the coefficient α

The factor α is determined to minimize the potential ϕ_{k+1} in the next step. The potential ϕ_{k+1} in the next step is

$$\begin{aligned}
\phi_{k+1} &= \frac{1}{2}(\mathbf{x}_{k+1}, \mathbf{A}\mathbf{x}_{k+1}) - (\mathbf{x}_{k+1}, \mathbf{f}) && (19)\\
&= \frac{1}{2}((\mathbf{x}_k + \alpha\mathbf{p}_k), \mathbf{A}(\mathbf{x}_k + \alpha\mathbf{p}_k)) - ((\mathbf{x}_k + \alpha\mathbf{p}_k), \mathbf{f}) && (20)\\
&= \left\{\frac{1}{2}(\mathbf{x}_k, \mathbf{A}\mathbf{x}_k) - (\mathbf{x}_k, \mathbf{f})\right\} + \alpha\left\{\frac{1}{2}(\mathbf{p}_k, \mathbf{A}\mathbf{x}_k) + \frac{1}{2}(\mathbf{x}_k, \mathbf{A}\mathbf{p}_k) - (\mathbf{p}_k, \mathbf{f})\right\} && (21)\\
&\quad +\alpha^2\left\{\frac{1}{2}(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)\right\} && (22)\\
&= \phi_k + \alpha\left(\mathbf{p}_k, \frac{\mathbf{A} + \mathbf{A}^T}{2}\mathbf{x}_k - \mathbf{f}\right) + \alpha^2\left\{\frac{1}{2}(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)\right\} && (23)
\end{aligned}$$

Since the matrix $\mathbf{A}$ was symmetric, it is $\frac{\mathbf{A}+\mathbf{A}^T}{2} = \mathbf{A}$. Accordingly

$$\phi_{k+1} = \phi_k + \alpha(\mathbf{p}_k, \mathbf{r}_k) + \alpha^2\left\{\frac{1}{2}(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)\right\} \tag{24}$$

When the potential ϕ_{k+1} takes the minimum value, the potential ϕ_{k+1} takes an extreme value

$$\frac{\partial \phi_{k+1}}{\partial \alpha} = 0 \tag{25}$$

Therefore, when calculating this

$$(\mathbf{p}_k, \mathbf{r}_k) = \alpha(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k) \tag{26}$$

Therefore, the coefficient α becomes as follows.

$$\alpha = \frac{(\mathbf{p}_k, \mathbf{r}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} \tag{27}$$

Furthermore, using the relational expression $(\mathbf{r}_i, \mathbf{r}_j) = 0 \quad (\quad i \neq j \quad)$ of orthogonality of residuals to be described later,

$$(\mathbf{p}_k, \mathbf{r}_k) = (\mathbf{r}_k - \beta_{k-1}\mathbf{p}_{k-1}, \mathbf{r}_k) = (\mathbf{r}_k, \mathbf{r}_k) \tag{28}$$

As a result, the coefficient α_k can be further expressed as follows.

$$\alpha_k = \frac{(\mathbf{p}_k, \mathbf{r}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} = \frac{(\mathbf{r}_k, \mathbf{r}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)} \tag{29}$$

Using these, the algorithm of the CG method is as follows

3.4 algorithm (CG method)

1. *Compute* $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$, $\mathbf{p}_0 = \mathbf{r}_0$
2. *For* $k = 0, 1, \ldots, m$ *Do* :
   (a) $\alpha_k = \frac{(\mathbf{r}_k, \mathbf{r}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)}$
   (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) $\beta_k = \frac{(\mathbf{r}_{k+1}, \mathbf{r}_{k+1})}{(\mathbf{r}_k, \mathbf{r}_k)}$ item $\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k$
3. *End Do*

4 Krylov subspace and CG method

The CG method is a type of Krylov subspace method. The Krylov subspace method is a method for finding an approximate solution that is closest to the solution in the Krylov subspace.

Below we prove that the CG method is Krylov subspace method However, for subspace $\bar{\mathcal{K}}_k, \tilde{\mathcal{K}}_k, \mathcal{K}_k$

$$\bar{\mathcal{K}}_k = span\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_{k-1}\} \tag{30}$$

$$\tilde{\mathcal{K}}_k = span\{\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_{k-1}\} \tag{31}$$

$$\mathcal{K}_k = span\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \cdots, \mathbf{A}^{k-1}\mathbf{r}_0\} \tag{32}$$

.

4.1 Proof of zzzgkhvqn

Prove using mathematical induction

- bf $k = 1$
  It is obvious from $\mathbf{r}_0 = \mathbf{p}_0$.

- bf $k > 1$
  Assume that it holds at k. In this case, it is checked whether or not it holds for k + 1.
  $\mathbf{p}_k = \alpha(\mathbf{r}_k + \beta\mathbf{p}_{k-1})$, $\mathbf{p}_k \in \bar{\mathcal{K}}_{k+1}$, $\mathbf{r}_k \in \tilde{\mathcal{K}}_{k+1}$.
  Also from the induction hypothesis $\mathbf{p}_{k-1} \in \bar{\mathcal{K}}_k = \tilde{\mathcal{K}}_k$ $\bar{\mathcal{K}}_{k+1} = \tilde{\mathcal{K}}_{k+1}$ holds.
  $\mathbf{r}_k = \mathbf{r}_{k-1} + \alpha\mathbf{A}\mathbf{p}_{k-1}$, $\mathbf{r}_k \in \tilde{\mathcal{K}}_{k+1}$. According to the induction hypothesis, since $\mathbf{r}_{k-1} \in \tilde{\mathcal{K}}_k = \mathcal{K}_k$, $\mathbf{p}_{k-1} \in \mathcal{K}_k$, $\mathbf{A}\mathbf{p}_{k-1} \in \mathcal{K}k + 1$, $\tilde{\mathcal{K}}_{k+1} = \mathcal{K}_{k+1}$ is established
  From the above we can say $\bar{\mathcal{K}}_{k+1} = \tilde{\mathcal{K}}_{k+1} = \mathcal{K}_{k+1}$.
  It can be seen that k + 1 also holds when assuming that the proposition is established in k.

From the above, the proposition was proved by mathematical induction.

By the way, given the solution $\mathbf{x}_k$ at the k-th iteration,

$$\mathbf{x}_k - \mathbf{x}_0 = \sum_{i=0}^{k-1} \Delta\mathbf{x}_i = \sum_{i=0}^{k-1} \alpha_i \mathbf{p}_i \tag{33}$$

Because it is

$$\mathbf{x}_k - \mathbf{x}_0 \in \mathcal{K}_{k+1} = span\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \cdots, \mathbf{A}^k\mathbf{r}_0\} \tag{34}$$

It can be seen that it is. Therefore, it can be seen that the CG method is a Krylov subspace method that searches for solutions in Krylov subspace. Below, it shows that the CG method always converges with iteration of n solution from the condition that the residual r is orthogonal.

4.2 Proof of zzzzjrxju

Prove using mathematical induction

- When $k = 1$

$$\begin{aligned}
(\mathbf{r}_1, \mathbf{p}_0) &= (\mathbf{r}_0 - \alpha\mathbf{A}\mathbf{p}_0, \mathbf{p}_0) = (\mathbf{r}_0, \mathbf{p}_0) - \frac{(\mathbf{r}_0, \mathbf{p}_0)}{(\mathbf{p}_0, \mathbf{A}\mathbf{p}_0)}(\mathbf{A}\mathbf{p}_0, \mathbf{p}_0) = 0 && (35)\\
(\mathbf{A}\mathbf{p}_1, \mathbf{p}_0) &= (\mathbf{A}(\mathbf{r}_1 - \beta\mathbf{p}_0), \mathbf{p}_0) = (\mathbf{A}\mathbf{r}_1, \mathbf{p}_0) - \frac{(\mathbf{r}_1, \mathbf{A}\mathbf{p}_0)}{(\mathbf{p}_0, \mathbf{A}\mathbf{p}_0)}(\mathbf{A}\mathbf{p}_0, \mathbf{p}_0) && (36)\\
&= (\mathbf{r}_1, \mathbf{A}^T\mathbf{p}_0) - (\mathbf{r}_1, \mathbf{A}\mathbf{p}_0) = 0 && (37)
\end{aligned}$$

  Here we used the condition $\mathbf{A}^T = \mathbf{A}$ that $\mathbf{A}$ is symmetric item When $k > 1$ Assuming that it holds for $0 \leq i < j \leq k$, in order to investigate whether $0 \leq i < j \leq k+1$ also holds here, you can investigate whether it is true in $j = k+1$. In this case, considering the two cases of $i = k$ and $i \leq k$ separately

  - When $i = k$

$$\begin{aligned}
(\mathbf{r}_{k+1}, \mathbf{p}_i) &= (\mathbf{r}_{k+1}, \mathbf{p}_k) = (\mathbf{r}_k - \alpha\mathbf{A}\mathbf{p}_k, \mathbf{p}_k) && (38)\\
&= (\mathbf{r}_k, \mathbf{p}_k) - \frac{(\mathbf{r}_k, \mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)}(\mathbf{A}\mathbf{p}_k, \mathbf{p}_k) = 0 && (39)\\
(\mathbf{A}\mathbf{p}_{k+1}, \mathbf{p}_i) &= (\mathbf{A}\mathbf{p}_{k+1}, \mathbf{p}_k) = (\mathbf{A}(\mathbf{r}_{k+1} - \beta\mathbf{p}_k), \mathbf{p}_k) && (40)\\
&= (\mathbf{A}\mathbf{r}_{k+1}, \mathbf{p}_k) - \frac{(\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)}(\mathbf{A}\mathbf{p}_k, \mathbf{p}_k) && (41)\\
&= (\mathbf{r}_{k+1}, \mathbf{A}^T\mathbf{p}_k) - (\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_k) = 0 && (42)
\end{aligned}$$

– When $i < k$

$$
\begin{aligned}
(\mathbf{r}_{k+1}, \mathbf{p}_i) &= (\mathbf{r}_k - \alpha \mathbf{A}\mathbf{p}_k, \mathbf{p}_i) && (43)\\
&= (\mathbf{r}_k, \mathbf{p}_i) - \frac{(\mathbf{r}_k, \mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{A}\mathbf{p}_k)}(\mathbf{A}\mathbf{p}_k, \mathbf{p}_i) = 0 && (44)\\
(\mathbf{A}\mathbf{p}_{k+1}, \mathbf{p}_i) &= (\mathbf{p}_{k+1}, \mathbf{A}\mathbf{p}_i) = (\mathbf{r}_{k+1} - \beta \mathbf{p}_k, \mathbf{A}\mathbf{p}_i) && (45)\\
&= (\mathbf{r}_{k+1}, \mathbf{A}\mathbf{p}_i) - \beta(\mathbf{p}_k, \mathbf{A}\mathbf{p}_i) && (46)\\
&= \frac{1}{\alpha}(\mathbf{r}_{k+1}, \mathbf{r}_i - \mathbf{r}_{i+1}) - \beta(\mathbf{p}_k, \mathbf{A}\mathbf{p}_i) && (47)\\
&= \frac{1}{\alpha}\left(\mathbf{r}_{k+1}, (\mathbf{p}_i + \beta \mathbf{p}_{i-1}) - (\mathbf{p}_{i+1} + \beta \mathbf{p}_i)\right) - \beta(\mathbf{p}_k, \mathbf{A}\mathbf{p}_i) = 0 && (48)
\end{aligned}
$$

- So we can see that it also holds for $0 \leq i < j \leq k+1$ since it is true even when $j = k+1$

The proposition is proved by using the mathematical induction method from the above

4.3 Proof of zzznwvtdc

Prove by mathematical induction begin itemize item When $k = 1$

$$(\mathbf{r}_0, \mathbf{r}_1) = (\mathbf{p}_0, \mathbf{r}_1) = 0 \tag{49}$$

item When $k > 1$ Assuming that it holds for $0 \leq i < j \leq k$, in order to investigate whether $0 \leq i < j \leq k+1$ also holds here, you can check whether it is true in $j = k+1$.

$$
\begin{aligned}
(\mathbf{r}_i, \mathbf{r}_{k+1}) &= (\mathbf{r}_i, \mathbf{r}_k - \alpha \mathbf{A}\mathbf{p}_k) = (\mathbf{r}_i, \mathbf{r}_k) - \alpha(\mathbf{A}\mathbf{p}_k, \mathbf{r}_i) && (50)\\
&= (\mathbf{r}_i, \mathbf{r}_k) - \alpha(\mathbf{A}\mathbf{p}_k, \mathbf{p}_i + \beta \mathbf{p}_{i-1}) = 0 && (51)
\end{aligned}
$$

Therefore, it can be seen that $0 \leq i < j \leq k+1$ also holds since it holds even in $j = k+1$ end itemize

The proposition is proved by using the mathematical induction method from the above

Here, it is found that the column $\mathbf{r}_i$ of the residual vector is a linear independent vector orthogonal to each other. Since the maximum value of the number of linearly independent vectors is the number of dimensions n of the vector, it can be seen that the number of columns of the residual vector does not become larger than the dimension number n of the vector. From this, in theory -

bf CG method always converges with the number of iterations of vector dimension n or less

It can be said that. However, this is only the theory of the theory and the CG method is strongly influenced by the rounding error, so on the calculator the CG method never converges per n iterations. Using the Chebyshev inequality, more accurate convergence conditions can be obtained.

5 CG method in complex matrix

If the coefficients are complex numbers, the coefficient matrix A is not a symmetric matrix but must be Hermetian Matrix. That is

$$A^* = A \tag{52}$$

In this case, the inner product (p, Ap) is a real number. Also, all eigenvalues are real numbers.

If the coefficient matrix is not a Hermitian matrix but a symmetric matrix, it can be solved using the Conjugate Orgonal Conjugate Gradient Method (COCG) method.

If not, use the CGNR method (Conjugate Non-Linear Method), etc.

The condition number of