Lecture 14

Derivation of conjugate gradient method (CG)

Now we'll derive a method, which uses the condition for nonzero diagonal entries of $P_m^T A P_m$ when A is pos. def. In this method, we'll also obtain recursions for residuals.

Consider the case, when A is sym. pos. def.

Search the approximate solution and the search direction in the form:

(1) $x_{j+1} = x_j + \alpha_j p_j$

(2) $p_{j+1} = r_{j+1} + \beta_j p_j$

Here α_j and β_j are unknown scalars to be found.

We'll use:

1) the orthogonality for residuals:

$(r_i, r_j) = 0, \; i \neq j$

2) the conjugacy for search directions:

$$(Ap_i, p_j) = \begin{cases} 0, & i \neq j \\ >0, & i = j \end{cases}$$

1) Let's find α_j using the first condition:

$(r_i, r_j) = 0, \; i \neq j$

Consider r_j and r_{j+1}

$(r_{j+1}, r_j) = 0$

$$\begin{cases} r_{j+1} = b - A x_{j+1} \\ r_j = b - A x_j \end{cases}$$

$$r_{j+1} - r_j = \cancel{b} - Ax_{j+1} - \cancel{b} + Ax_j =$$

$$= -A\underbrace{x_{j+1}}_{x_j + \alpha_j p_j} + Ax_j = \cancel{-Ax_j} - \alpha_j A p_j \cancel{+Ax_j} = -\alpha_j A p_j$$

$\Rightarrow$

$$\boxed{r_{j+1} = r_j - \alpha_j A p_j} \quad (3)$$ recursion for residuals

$\Rightarrow$

$$(r_{j+1}, r_j) = (r_j - \alpha_j A p_j, r_j) = 0$$

$$(r_j, r_j) - \alpha_j (Ap_j, r_j) = 0 \Rightarrow$$

$$\boxed{\alpha_j = \frac{(r_j, r_j)}{(Ap_j, r_j)}} \quad (4)$$ we'll modify this formula later

2) Now let's find β_j using the second condition:

$$(Ap_i, p_j) = \begin{cases} 0, & i \neq j \\ > 0, & i = j \end{cases}$$

Consider p_j and p_{j+1}:

$$p_{j+1} = r_{j+1} + \beta_j p_j$$

$$(Ap_j, p_{j+1}) = 0$$

$$(Ap_j, r_{j+1} + \beta_j p_j) = 0$$

$$\underbrace{(Ap_j, r_{j+1})} + \beta_j (Ap_j, p_j) = 0 \quad (*)$$

$(Ap_j, r_{j+1}) =$

$$r_{j+1} = r_j - \alpha_j A p_j \Rightarrow Ap_j = -\frac{1}{\alpha_j}(r_{j+1} - r_j)$$

$$\ominus \left(-\frac{1}{\alpha_j}(z_{j+1}-z_j),\ z_{j+1}\right)=$$

$$=-\frac{1}{\alpha_j}(z_{j+1},z_{j+1})+\frac{1}{\alpha_j}\underbrace{(z_j,z_{j+1})}_{=0} \Rightarrow$$

$$(Ap_j,\ z_{j+1})=-\frac{1}{\alpha_j}(z_{j+1},z_{j+1}) \ominus$$

$$\alpha_j=\frac{(z_j,z_j)}{(Ap_j,z_j)}$$

$$\ominus -\frac{(Ap_j,z_j)}{(z_j,z_j)}(z_{j+1},z_{j+1})$$

Thus, in $(*)$ we have

$$-\frac{(Ap_j,z_j)}{(z_j,z_j)}(z_{j+1},z_{j+1})+\beta_j(Ap_j,p_j)=0 \Rightarrow$$

$$\beta_j=\frac{(Ap_j,z_j)(z_{j+1},z_{j+1})}{\underbrace{(z_j,z_j)}_{\neq 0}\underbrace{(Ap_j,p_j)}_{>0 \text{ as } A \text{ is pos. def.}}}$$

As $p_{j+1}=z_{j+1}+\beta_j p_j$, then

$$z_{j+1}=p_{j+1}-\beta_j p_j$$

$$z_j=p_j-\beta_{j-1}p_{j-1}$$

Consider $(Ap_j,z_j)=(Ap_j,\ p_j-\beta_{j-1}p_{j-1})=$

$$=(Ap_j,p_j)-\beta_{j-1}\underbrace{(Ap_j,p_{j-1})}_{=0}=(Ap_j,p_j)$$

Hence, $$\beta_j=\frac{\cancel{(Ap_j,p_j)}(z_{j+1},z_{j+1})}{(z_j,z_j)\cancel{(Ap_j,p_j)}}=\frac{(z_{j+1},z_{j+1})}{(z_j,z_j)}$$

Thus, $$\boxed{\beta_j = \frac{(r_{j+1}, r_{j+1})}{(r_j, r_j)}} \quad (5)$$

and $$\boxed{\alpha_j = \frac{r_j, r_j}{Ap_j, p_j}} \quad (4 \text{ new})$$

We have obtained 5 recursions:

- for approximate solution and residual

1) $\alpha_j = \dfrac{(r_j, r_j)}{(Ap_j, p_j)}$ (4)

2) $x_{j+1} = x_j + \alpha_j p_j$ (1)

3) $r_{j+1} = r_j - \alpha_j A p_j$ (3)

- for search direction

4) $\beta_j = \dfrac{(r_{j+1}, r_{j+1})}{(r_j, r_j)}$ (5)

5) $p_{j+1} = r_{j+1} + \beta_j p_j$ (2)

These five recursions give conjugate gradient method (CG)

<u>Conjugate gradient method (CG)</u>

A is sym. pos. def.

1. Take initial guess x_0, compute initial residual $r_0 = b - Ax_0$; $p_0 := r_0$
2. Start iterative process for $j = 0, 1, 2, \ldots$ Do until convergence

1) $\alpha_j = \dfrac{(r_j, r_j)}{(Ap_j, p_j)}$

2) $x_{j+1} = x_j + \alpha_j p_j$ ← check convergence of x_{j+1}

3) $r_{j+1} = r_j - \alpha_j A p_j$ ← check residual norm $\| r_{j+1} \|_2 < \varepsilon$

If not converged,

4) $\beta_j = \dfrac{(r_{j+1}, r_{j+1})}{(r_j, r_j)}$

5) $p_{j+1} = r_{j+1} + \beta_j p_j$

As it was in Classic Lanczos and D-Lanczos, in CG method the residuals r_j are orthogonal and search directions p_j are conjugate. Vectors p_j in CG are called <u>conjugate gradients</u>

(T) In CG-method A is sym. pos. def. and

1) residuals r_j are orthogonal:

$$(r_i, r_j) = \begin{cases} 0, & i \neq j \\ > 0, & i = j \end{cases}$$

2) search directions p_j are A-conjugate:

$$(A p_i, p_j) = \begin{cases} 0, & i \neq j \\ > 0, & i = j \end{cases}$$

(T) For $\mathrm{cond}(A) \simeq 1$ CG converges very fast