Lecture 16

Thursday, 3 March 2022 1:33 PM

CG method

I. $\{ \underline{d}^{(1)}, \underline{d}^{(2)}, \dots, \underline{d}^{(n)} \}$ mutually orthogonal; one step in one direction.

$$\boxed{\underline{x}^{(k+1)} = \underline{x}^{(k)} + \alpha_k \underline{d}^{(k)}};$$

$$\boxed{\alpha_k = -\frac{(\underline{d}^{(k)})^T \underline{e}^{(k)}}{(\underline{d}^{(k)})^T \underline{d}^{(k)}}}\ ??$$

Useless.

II. Change to Q orthogonal search directions.

$$(\underline{d}^{(j)})^T Q \, \underline{d}^{(k)} = 0$$

$$\boxed{\alpha_k = -\frac{(\underline{d}^{(k)})^T \underline{g}^{(k)}}{(\underline{d}^{(k)})^T Q \, \underline{d}^{(k)}}}$$

Can be computed.

III. 'Q' orthogonal directions $\{\underline{d}^{(j)}\}_{j=1}^{n}$ $\Rightarrow$ Convergence after '(n+1)' steps.

$\rightarrow$ Q orthogonal directions are l.i. if Q is SPD.
$\hookrightarrow$ (Proved in last class)

$$\underline{e}^{(1)} = \sum_{j=1}^{n} \delta_j \, \underline{d}^{(j)}$$

$\times$ by $(\underline{d}^{(k)})^T Q$. $\rightarrow$

$$(\underline{d}^{(k)})^T Q \, \underline{e}^{(1)} = \sum_{j=1}^{n} \delta_j (\underline{d}^{(k)})^T Q \, \underline{d}^{(j)}$$

$$= \delta_k (\underline{d}^{(k)})^T Q \, \underline{d}^{(k)}$$

$$\delta_k = \frac{(\underline{d}^{(k)})^T Q \, \underline{e}^{(1)}}{(\underline{d}^{(k)})^T Q \, \underline{d}^{(k)}} \checkmark$$

$$= \frac{(\underline{d}^{(k)})^T Q \overbrace{\left[e^{(1)} + \underbrace{\sum_{j=1}^{k-1} \alpha_j \underline{d}^{(j)}}_{\text{" } 0} \right]}^{\underline{e}^{(k)}}}{(\underline{d}^{(k)})^T Q \, \underline{d}^{(k)}}$$

Since $(\underline{d}^{(k)})^T Q \, \underline{d}^{(j)} = 0$.

$$x^{(k)} - x^{(*)} = x^{(1)} - x^{(*)} + \dots$$

$$\underline{e}^{(k)} = e^{(1)} + \sum_{j=1}^{k-1} \alpha_j \underline{d}^{(j)}$$

$$x^{(k)} = x^{(k-1)} + \alpha_{k-1} \underline{d}^{(k-1)}$$

$$\underline{r}^{(k)} = Q e^{(k)}$$

$$= \frac{\ldots}{(d^{(k)})^T \, Q \, d^{(k)}}$$

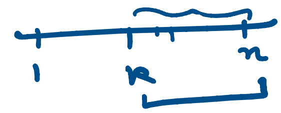
Since $(d^{(k)})^T Q d^{(j)} = 0$.

$$= \frac{(\underline{d}^{(k)})^T \, Q \, e^{(k)}}{(\underline{d}^{(k)})^T \, Q \, \underline{d}^{(k)}} = \frac{(\underline{d}^{(k)})^T \, \underline{r}^{(k)}}{(d^{(k)})^T \, Q \, d^{(k)}}$$

$$\delta_k = -\alpha_k.$$

$$\underline{e}^{(k)} = e^{(1)} + \sum_{j=1}^{k-1} \alpha_j d^{(j)}$$

$$= \sum_{j=1}^{n} \delta_j \underline{d}^{(j)} - \sum_{j=1}^{k-1} \delta_j d^{(j)}$$



$$= \sum_{j=k}^{n} \delta_j \underline{d}^{(j)}$$

$$\vdots$$

$$\underline{e}^{(n+1)} = 0.$$

$\underline{\text{IV}}$. How to construct Q <u>orthogonal search directions</u>? [Gram-Schmidt process].

$\{\underline{d}^{(1)}, \dots, \underline{d}^{(n)}\} \to$ l.i.

<u>Recall Gram-Schmidt</u>: $\{\underline{v}_1, \dots, \underline{v}_n\}$ l.i.

$\{\underline{e}_1, \dots, \underline{e}_n\}$ ons

$$\underline{u}_1 = \underline{v}_1, \quad \underline{e}_1 = \frac{\underline{u}_1}{\|u_1\|}$$

$$\underline{u}_2 = \underline{v}_2 - \langle \underline{v}_2, \underline{u}_1 \rangle \underline{u}_1 \qquad e_2 = \frac{\underline{u}_2}{\|u_2\|}$$

not needed

$$\underline{u}_2 = \underline{v}_2 - \frac{\langle \underline{v}_2, \underline{u}_1 \rangle}{\langle \underline{u}_1, \underline{u}_1 \rangle} \underline{u}_1 \qquad e_2 = \frac{\underline{u}_2}{\|\underline{u}_2\|}$$

$$\vdots$$

$$\underline{u}_n = \underline{v}_n - \frac{\langle \underline{v}_n, \underline{u}_1 \rangle}{\langle \underline{u}_1, \underline{u}_1 \rangle} \underline{u}_1 - \cdots \frac{\langle \underline{v}_n, \underline{u}_{n-1} \rangle}{\langle \underline{u}_{n-1}, \underline{u}_{n-1} \rangle} \underline{u}_{n-1}; \qquad e_n = \frac{\underline{u}_n}{\|\underline{u}_n\|}$$

Modify Gram-Schmidt for Q_orthogonality

Let $\{\underline{u}^{(i)}\}_{i=1}^{n}$ be n- l.i. directions.

Aim: To construct $\{\underline{d}^{(i)}\}_{i=1}^{n}$ that are 'Q' orthogonal.

in such a way that it is

[$d^{(k)}$ is constructed
Q orthogonal to $(\underline{d}^{j})_{j=1}^{k-1}$

$$\underline{d}^{(1)} = u^{(1)}.$$

$$\underline{d}^{(k)} = \underline{u}^{(k)} + \sum_{j=1}^{k-1} \underline{\underline{\beta_{kj}}}\, \underline{d}^{(j)}$$

β_{kj}'s are to be chosen such that $\{\underline{d}^{(j)}\}_{j=1}^{k-1}$ are Q - orthogonal to $\underline{d}^{(k)}$.

[Try with a single value of j; $j=1$]

For $\boxed{j = 1, \ldots, k-1}$;

$$(Q\,\underline{d}^{(j)})^T \underline{d}^{(k)} = (\underline{d}^{(j)})^T Q \left(\underline{u}^{(k)} + \sum_{j=1}^{k-1} \beta_{kj}\, \underline{d}^{(j)} \right)$$

$$0 = (\underline{d}^{(j)})^T Q\, \underline{u}^{(k)} + \beta_{kj} (\underline{d}^{(j)})^T Q\, \underline{d}^{(j)}$$

$$\Rightarrow \boxed{\beta_{kj} = - \frac{(\underline{u}^{(k)})^T Q\, \underline{d}^{(j)}}{(\underline{d}^{j})^T Q\, \underline{d}^{(j)}}}$$

$$\boxed{\underline{d}^{(k)} = \underline{u}^{(k)} + \sum_{j=1}^{k-1} \beta_{kj}\, \underline{d}^{(j)}}$$

Drawbacks.

- Search directions need to be stored
- Linear combinations.

V. [Choice of $\{u^{(i)}\}_{i=1}^{n}$ (l.i) such that the drawbacks above are sorted out.

Choose $\underline{u}^{(i)} = \underline{r}^{(i)}$ $\rightarrow$ • Can we do this?

Qus: Are $\{\underline{r}^{(k)}\}_{k=1}^{n}$ l.i.?

Do they simplify computation δ_k $\underline{d}^{(k)}$?

Properties of residuals:

$$\boxed{\underline{r}^{(k)} = Q e^{(k)}}$$

(P1) $\underline{r}^{(k)}$ is orthogonal to $\boxed{\{\underline{d}^{(i)}\}_{i=1}^{k-1}}$

Pf: $\underline{e}^{(k)} = \sum_{j=k}^{n} \delta_j\, \underline{d}^{(j)}$.

$i = 1, \dots, k-1$

$$\underbrace{(d^{(i)})^T Q\, \underline{e}^{(k)}} = \sum_{j=k}^{n} \delta_j\, (d^{(i)})^T Q\, d^{(j)}$$

$$(d^{(i)})^T\, \underline{r}^{(k)} = 0 . \qquad i = 1, \dots, k-1.$$

(P2) Claim $\underline{r}^{(k+1)} = \underline{r}^{(k)} + \alpha_k\, Q\, \underline{d}^{(k)}$.

$$\boxed{x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}}$$

Pf. $\underline{r}^{(k+1)} = Q\, \underline{e}^{(k+1)}$

Pf. $r^{(k+1)} = Q e^{(k+1)}$

$= Q\left(e^{(k)} + \alpha_k d^{(k)}\right)$

$$\boxed{r^{(k+1)} = r^{(k)} + \alpha_k Q d^{(k)}}$$

(P3) $$\beta_{kj} = -\frac{(r^{(k)})^T Q d^{(j)}}{(d^{(j)})^T Q d^{(j)}}$$ [from derivation of β_{kj}'s with $u^{(k)} = r^{(k)}$].

(P4) $\boxed{\{r^{(j)}\}_{j=1}^{n} \text{ are orthogonal}}$: $\rightarrow (r^{(k)})^T r^{(j)} = 0$, $k \neq j$.

HW

Hint: use (P1)

$$d^{(k)} = u^{(k)} + \sum_{j=1}^{k-1} \beta_{kj} d^{(j)}. \quad \| \cdot \times (r^{(j)})^T$$

(P5) Most β_{kj}'s are zeroes when $u^{(j)} = r^{(j)}$.

$$d^{(k)} = r^{(k)} + \sum_{j=1}^{k-1} \beta_{k,j} d^{(j)}$$

$$= r^{(k)} + \beta_{k,1} d^{(1)} + \dots + \beta_{k,k-1} d^{(k-1)}$$

$$\boxed{\beta_{k,j} = -\frac{(r^{(k)})^T Q d^{(j)}}{(d^{(j)})^T Q d^{(j)}}}$$ ↳ (P3)

$$r^{(j+1)} = r^{(j)} + \alpha_j Q d^{(j)}$$ $\boxed{\text{from (P2)}}$

$\times$ by $(r^{(k)})^T$

$$(r^{(k)})^T r^{(j+1)} = (r^{(k)})^T r^{(j)} + \alpha_j \cdot (r^{(k)})^T Q d^{(j)}$$

$$\alpha_j (g^{(k)})^T Q d^{(j)} = (g^{(k)})^T g^{(j+1)} - (g^{(k)})^T g^{(j)}$$

$$\boxed{(g^{(k)})^T Q d^{(j)} = \begin{cases} -\frac{1}{\alpha_k} (g^{(k)})^T g^{(k)} & j = k \\ \checkmark \frac{1}{\alpha_{k-1}} (g^{(k)})^T g^{(k)} \longleftarrow & j+1 = k \\ 0 & \text{Otherwise} \end{cases}}$$

$$\beta_{kj} = - \frac{(g^{(k)})^T Q d^{(j)}}{(d^{(j)})^T Q d^{(j)}}$$

$$\beta_k = \beta_{k,k-1} = -\frac{1}{\alpha_{k-1}} \frac{(g^{(k)})^T g^{(k)}}{(d^{(k-1)})^T \underbrace{Q d^{(k-1)}}}$$

$$\boxed{\beta_{k1}, \beta_{k2} \cdots, \beta_{k,k-2} = 0}$$

$$\alpha_{k-1} = - \frac{(d^{(k-1)})^T g^{(k-1)}}{\underbrace{(d^{(k-1)})^T Q d^{(k-1)}}}$$

$$\boxed{\beta_k = \frac{(g^{(k)})^T g^{(k)}}{(d^{(k-1)})^T g^{(k-1)}}}.$$

HW $(d^{(k-1)})^T g^{(k-1)} = (g^{(k-1)})^T g^{(k-1)}$.

Hint. $d^{(k-1)} = g^{(k-1)} + \sum_{j=1}^{k-2} \beta_{k-1,j} d^{(j)}$

$$\boxed{\beta_k = \frac{(g^{(k)})^T g^{(k)}}{(g^{(k-1)})^T g^{(k-1)}}}.$$

$$\boxed{\beta_k = \frac{(\underline{r}^{(k)})^T \underline{r}^{(k)}}{(\underline{r}^{(k-1)})^T \underline{r}^{(k-1)}}}$$

eg.