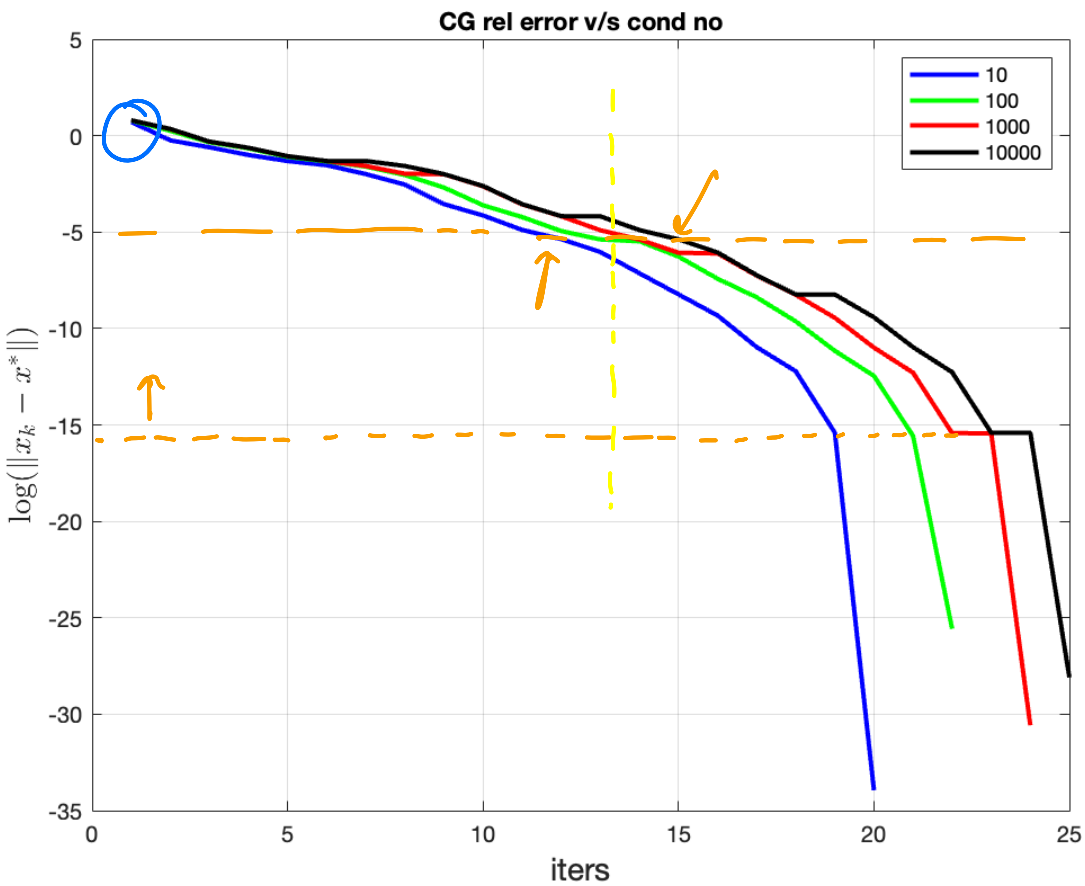

Observations : Condn no plays a big role in speed of convergence.

Preconditioned CG.

$$\phi(x) = \frac{1}{2} x^T A x - b^T x, \quad \text{Min this}$$

Same as solving $\underline{Ax = b}$.

$$\underline{PA}x = \underline{P}b$$

Too expensive $\quad P = A^{-1} \Rightarrow Ix = Pb$

preconditioning A with P.

$[\,\mathrm{Cond}(AB) \neq \mathrm{Con}(A)\,\mathrm{cond}(B).\,]$

Common trick: $LL^T A x = LL^T b.$

where L is full rank.

$$L^T A x = L^T b$$

$$I = L L^{-1}$$

$$\Rightarrow \underbrace{L^T A L}_{\hat{A}} \underbrace{L^{-1} x}_{= \hat{x}} = \underbrace{L^T b}_{= \hat{b}} \longrightarrow \text{is sym pos def?}$$

$$z^T L^T A L z = (Lz)^T A (Lz) > 0 \quad \checkmark$$

① $\hookrightarrow$ Solve $\boxed{\hat{A}\hat{x} = \hat{b}}$ by CG machinery.

once done $L^{-1}x = \hat{x} \Rightarrow \underline{x = L\hat{x}}.$

efficiency $\times$.

$\boxed{Ax = b}$

② Something more efficient.

$\hookrightarrow$ $\underline{\hat{r}_k} = \hat{A}\hat{x}_k - \hat{b} \xrightarrow{?} r_k$ $\quad \hat{x} = L^{-1}x.$

$$= (L^T A L)(L^{-1}x_k) - L^T b$$

$$= L^T [A x_k - b] = \underline{L^T r_k} \quad \text{— ⓐ}$$

$\hookrightarrow$ $\alpha_k = r_k^T r_k$ $\quad$ Hat: $\hat{p}_k^T \hat{A} \hat{p}_k$

$$\underbrace{p_k^T A p_k}$$

$$\hat{p}_k^T (L^T A L) \hat{p}_k \qquad \boxed{\hat{p}_k = L^{-1} p_k}$$

$$p_k^T L^{-T} (L^T A L) L^{-1} p_k = p_k^T A p_k$$

$$\hat{\alpha}_k = \frac{\hat{r}_k^T \hat{r}_k}{\hat{p}_k^T \hat{A} \hat{p}_k} = \frac{r_k^T L L^T r_k}{p_k^T A p_k} \quad \text{— (6)}$$

$$\boxed{\hat{\beta}_{k+1} = \frac{\hat{r}_{k+1}^T \hat{r}_{k+1}}{\hat{r}_k^T \hat{r}_k}} = \frac{r_{k+1}^T L L^T r_{k+1}}{r_k^T L L^T r_k}$$

$LL^T = M^{-1}$ defn.

$$M^{-1} r_k \triangleq y_k.$$

$$\hat{\alpha}_k = \frac{r_k^T y_k}{p_k^T A p_k}$$

$$\hat{\beta}_{k+1} = \frac{r_{k+1}^T y_{k+1}}{r_k^T y_k}$$

$$p_k = -r_k + \beta_k p_{k-1}$$

$$\hat{p}_k = -\hat{r}_k + \hat{\beta}_k \hat{p}_{k-1}$$

$$\Rightarrow \quad L^{-1} p_k = -L^T r_k + \hat{\beta}_k L^{-1} p_{k-1}$$

$$p_k = -\underset{\sim}{LL^T} r_k + \hat{\beta}_k p_{k-1}$$

$$p_k = -y_k + \hat{\beta}_k p_{k-1}$$

$$\hat{p}_0 = -\hat{r}_0 = -L^T r_0 = L^{-1} p_0$$

$$p_0 = -LL^T r_0 = -y_0$$

$$\hat{x}_{k+1} = \hat{x}_k + \hat{\alpha}_k \hat{p}_k$$

$$\cancel{L^{-1}} x_{k+1} = \cancel{L^{-1}} x_k + \hat{\alpha}_k \cancel{L^{-1}} p_k$$

I must store y in addn.

CG iters $\rightarrow$ They update r_k, p_k

Solve lin sys $M^{-1} r = y$

Extra cost.

↳ Incomplete Cholesky decomp.

$A = CC^T$ exact.

$A \approx \tilde{C}\tilde{C}^T$

eigvalues of $(L^T A L)$

$L^T \tilde{C} \tilde{C}^T L$

$\approx I.$

$(L^T)^{-1} = C$

$Ax_k - b$

$p_k = -r_k + \beta_k p_{k-1}$

Nonlinear CG.

$\phi(x) = \frac{1}{2} x^T A x - b^T x \longleftrightarrow Ax = b$

sym pos def.

↳ Can ϕ be general convex fn?

↳ Can ϕ be a nonlinear fn?

Yes! $f(x)$.

↳ 2 changes needed linear CG.

① step length: no exact expr for α_k.

line search for α_k.

② Residual : replace by $\nabla f(x) = r(x)$.

→ Start with $p_0 = -\nabla f_0$

→ $$\beta_{k+1} = \frac{\nabla f_{k+1}^T \nabla f_{k+1}}{\nabla f_k^T \nabla f_k}$$

→ $$p_{k+1} = -\nabla f_{k+1} + \beta_{k+1} p_k$$

→ α_k from line search.

Fletcher Reeves method