①

$$r_0 = b - Ax_0$$

$$d_0 = M^{-1} r_0$$

$$\alpha_i = \frac{r_i^t M^{-1} r_i}{d_i^t A d_i}$$

$$x_{i+1} = x_i + \alpha_i d_i$$

$$r_{i+1} = r_i - \alpha_i A d_i$$

$$\beta_{i+1} = \frac{r_{i+1}^t M^{-1} r_{i+1}}{r_i^t M^{-1} r_i}$$

$$d_{i+1} = M^{-1} r_{i+1} + \beta_{i+1} d_i$$

$\rightarrow$ setting $\tilde{M} = \gamma M$.

$$\tilde{r}_0 = b - A\tilde{x}_0$$

$$\tilde{d}_0 = \tilde{M}^{-1} \tilde{r}_0$$

$$\tilde{\alpha}_i = \frac{\tilde{r}_i^t \tilde{M}^{-1} \tilde{r}_i}{\tilde{d}_i^t A \tilde{d}_i}$$

$$\tilde{x}_{i+1} = \tilde{x}_i + \tilde{\alpha}_i \tilde{d}_i$$

$$\tilde{r}_{i+1} = \tilde{r}_i - \tilde{\alpha}_i A \tilde{d}_i$$

$$\tilde{\beta}_{i+1} = \frac{\tilde{r}_{i+1}^t \tilde{M}^{-1} \tilde{r}_{i+1}}{\tilde{r}_i^t \tilde{M}^{-1} \tilde{r}_i}$$

$$\tilde{d}_{i+1} = \tilde{M}^{-1} \tilde{r}_{i+1} + \tilde{\beta}_{i+1} \tilde{d}_i$$

$\leftarrow$

WTS

$\tilde{x}_i = x_i$, $\tilde{r}_i = r_i$

$\tilde{\alpha}_i = \gamma \alpha_i$, $\tilde{d}_i = \frac{1}{\gamma} d_i$

$\tilde{\beta}_i = \beta_i$

$\because$

Tl $i = 1$

$$\tilde{r}_0 = b - A\tilde{x}_0$$

$$\tilde{d}_0 = \tilde{M}^{-1} \tilde{r}_0 = (1/\gamma) d_0$$

$$\tilde{\alpha}_0 = \frac{\tilde{r}_0^t \tilde{M}^{-1} \tilde{r}_0}{\tilde{d}_0^t A \tilde{d}_0} = \gamma \alpha_0$$

$$\tilde{x}_1 = \tilde{x}_0 + \tilde{\alpha}_0 \tilde{d}_0 = x_0 + \alpha_0 d_0 = x_1$$

$$\tilde{r}_1 = \tilde{r}_0 - \tilde{\alpha}_0 A \tilde{d}_0 = r_0 - \alpha_0 A d_0 = r_1$$

$$\tilde{\beta}_1 = \frac{\tilde{r}_1{}^t \tilde{M}^{-1} \tilde{r}_1}{\tilde{r}_0{}^t \tilde{M}^{-1} \tilde{r}_0} = \frac{r_1{}^t M^{-1} r_1}{r_0{}^t M^{-1} r_0} = \beta_1$$

$$\tilde{d}_1 = \tilde{M}^{-1}\tilde{r}_1 + \tilde{\beta}_1 \tilde{d}_0$$

$$= \frac{1}{r}(M^{-1}r_1) + \beta_1 \cdot \frac{1}{r} d_1 = \frac{1}{r} d_1 .$$

iii) If true for $1, 2, \cdots, i$

$$\tilde{x}_{i+1} = \tilde{x}_i + \tilde{\alpha}_i \tilde{d}_i = x_i + r\alpha_i \cdot \frac{1}{r} d_i = x_i + \alpha_i d_i = x_{i+1}$$

$$\tilde{r}_{i+1} = \tilde{r}_i - \tilde{\alpha}_i A \tilde{d}_i = r_i - r\alpha_i \cdot A \cdot \frac{1}{r} d_i = r_i - \alpha_i A d_i = r_{i+1}$$

$$\tilde{\beta}_{i+1} = \frac{\tilde{r}_{i+1}{}^t \tilde{M}^{-1} \tilde{r}_{i+1}}{\tilde{r}_i{}^t \tilde{M}^{-1} \tilde{r}_i} = \frac{r_{i+1} \cdot M^{-1} \cdot \frac{1}{r} \cdot r_{i+1}}{r_i M^{-1} \cdot \frac{1}{r} \cdot r_i} = \beta_{i+1}$$

$$\tilde{d}_{i+1} = \tilde{M}^{-1}\tilde{r}_{i+1} + \tilde{\beta}_{i+1} \tilde{d}_i = \frac{1}{r} M^{-1} r_{i+1} + \beta_{i+1} \cdot \frac{1}{r} d_i = \frac{1}{r} d_{i+1}$$

$$\tilde{\alpha}_{i+1} = \frac{\tilde{r}_{i+1}{}^t \tilde{M}^{-1} \tilde{r}_{i+1}}{\tilde{d}_{i+1}{}^t A \tilde{d}_{i+1}} = \frac{r_{i+1} \cdot \frac{1}{r} M^{-1} \cdot r_{i+1}}{\frac{1}{r} d_{i+1}{}^t A \frac{1}{r} d_{i+1}} = r \cdot \alpha_{i+1}$$

$\blacksquare$

②

(a) For A is sym. p.d., it is diagonalizable.

Then, $\exists$ a set of orthonormal vectors, which is also an eigvec of A.

See x_0 as $\sum_i a_i v_i$,

$\hookrightarrow -Ax_0 + b = -\sum_i \lambda_i a_i v_i + b$

$$\hookrightarrow \quad x_1 = \sum_i a_i v_i - \frac{1}{\lambda_1} \sum_i \lambda_i a_i v_i + \frac{b}{\lambda_1}$$

$$= \sum_{\substack{i \text{ s.t.} \\ \lambda_i \neq \lambda_1}} \tilde{a}_i v_i + \frac{b}{\lambda_1} .$$

repeating this step., we get

$$x_2 = \sum_{\substack{i \text{ s.t.} \\ \lambda_i \neq \lambda_1, \lambda_2}} \tilde{\tilde{a}}_i v_i + \underbrace{\frac{b}{\lambda_1} + \frac{1}{\lambda_2}\left(b - \frac{Ab}{\lambda_1}\right)} .$$

$$A^{-1}\left[- \prod_{i=1}^{2} \left(I - \frac{A}{\lambda_i} \right) + I \right] b$$

$$x_3 = \sum_{\substack{i \text{ s.t.} \\ \lambda_i \neq \lambda_1, \lambda_2, \lambda_3}} \tilde{\tilde{a}}_i v_i + A^{-1}\left[- \prod_{i=1}^{3} \left(I - \frac{A}{\lambda_i} \right) + I \right] b$$

WTS $x_k \, (k \leq d) = \underbrace{\sum_{\substack{i \text{ s.t.} \\ \lambda_i \neq \lambda_1, \cdots, \lambda_k}} \hat{a}_i v_i}_{①} + \underbrace{A^{-1}\left[- \prod_{i=1}^{k} \left(I - \frac{A}{\lambda_i} \right) + I \right] b}_{②} .$

$\because$ ① term is quite trivial. For the ②, updated form is of

② $+ \frac{1}{\lambda_{k+1}}(b - A \cdot ②)$, we use induction. (k=2 case is previously proved)

$$A^{-1}\left[- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I \right] + \frac{1}{\lambda_{n+1}} \left(b - A \cdot A^{-1} \left[- \prod_{i=1}^{3} \left(I - \frac{A}{\lambda_i} \right) + I \right] \right)$$

$$= A^{-1}\left[- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I \right] b + \frac{1}{\lambda_{n+1}} \left(b - \left[- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I \right] b \right)$$

$$= A^{-1} \left(- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I + \frac{1}{\lambda_{n+1}} \left(I - \left[- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I \right] \right) \right) b$$

$$= A^{-1} \left(- \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + \frac{1}{\lambda_{n+1}} \prod_{i=1}^{n} \left(I - \frac{A}{\lambda_i} \right) + I \right) b$$

$$= \underline{A^{-1}\left[- \prod_{i=1}^{n+1} \left(I - \frac{A}{\lambda_i} \right) + I \right] b} \; //$$

$\therefore$ Iteration stops till d.

$$\leadsto x_d = A^{-1}\left[- \prod_{i=1}^{d} (I - \frac{A}{\lambda_i}) + I \right] b.$$

For A is diagonalizable, the minimal polynomial is in form of

$$m_A(x) = \prod_{j=1}^{d} (x - \lambda_j)^{e_j} \quad (0 \le e_j \le 1)$$

$$x_d = A^{-1}(- \prod_{i=1}^{d} (I - \frac{A}{\lambda_i}) + I) b = A^{-1}(- \underbrace{m_A(A)}_{}^{0} \cdot p(A) + I) b$$

$$= A^{-1} \cdot I \cdot b = A^{-1} b$$

$\therefore x_d$ is solution of such system.

(b) To speed up the convergence, we find the value $r_i^t r_i$, or $r_i^t A r_i$. For $r_i = b - Ax_i$, selecting smaller r_i is satisfactory

$$r_{i+1}^t r_{i+1} = [b - A(x_i + \lambda_i^{-1}(b - Ax_i))]^t A [b - A(x_i + \lambda_i^{-1}(b - Ax_i))]$$

WT find λ_i, s.t. minimizes $\| b - A(x_i + \lambda_i^{-1}(b - Ax_i)) \|_2$.

$\frac{r_{i+1}^t r_{i+1}}{r_i^t r_i}$ minimize. $r_i^t r_i$ fixed.

For A, b, x_i is known, it can be shown that this is function on λ_i

It's better to choose smaller λ_i. for large λ_i would make $\lambda_i^{-1} \simeq 0$, which make not much difference on $x_i + \lambda_i^{-1}(b - Ax_i)$, compared with x_i.

(c)

For $x_{i+1} = x_i + \lambda_i^{-1} r_i$.

↳ If eigval becomes small, floating point error would likely to occur

(d) To prevent floating pt errors we choose eigval from middle range. For huge eigval or small eigval can cause floating pt error, they are chosen at the latter part.

↳ or, not calculating them could also give finer result.

In any ways, try to avoid calculating the small or large eigval / related calculation.