- (SD example) We consider the steepest decent method applied to the following data: For a real $a \gg 1$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x_0 = \begin{pmatrix} a \\ 1 \end{pmatrix}$$

(a) Show that the even iterates are

$$x^{(2m)} = \left(\frac{a-1}{a+1}\right)^{2m} x_0, \quad m \in \mathbb{N},$$

and find a formula for the odd iterates.
Compute the $\alpha^{(k)}$ and the residuals $r^{(k)}$, $k \in \mathbb{N}$.

(b) Show that subsequent search directions are 'almost' parallel in the metric defined by A: Use the formula

$$\cos(\phi) = \frac{\langle x, y\rangle_A}{\|x\|_A \|y\|_A}$$

to measure the angle ϕ between to subsequent search directions and check that ϕ tends to π as $a \to \infty$.

Remark: In contrast, the search directions in the CG method are orthogonal with respect to $\langle \cdot, \cdot \rangle_A$ which results in far better convergence properties.

Soln

We know that

$$\begin{aligned} x^{(k)} &= x^{(k-1)} + \alpha^{(k-1)} d^{(k-1)} \\ &= x^{(k-1)} + \alpha^{(k-1)} r^{(k-1)} \end{aligned}$$

and

$$\alpha^{(k-1)} = \frac{\|r^{(k-1)}\|_2^2}{\|r^{(k-1)}\|_A^2}$$

Now we proceed to obtain an expression for $r^{(k-1)}$

$$r^{(0)} = b - Ax^{(0)} = -Ax^{(0)} \qquad (\text{Since } b = (0,0)^T)$$

$$\begin{aligned} \|r^{(0)}\|_2^2 &= (-Ax^{(0)})^T(-Ax^{(0)}) \\ &= x^{(0)T} A^T A x^{(0)} \\ &= x^{(0)T} A^2 x^{(0)} \qquad (\text{Since } A \text{ is diagonal}) \end{aligned}$$

$$\|r^{(0)}\|_A^2 = \langle r^{(0)}, Ar^{(0)} \rangle = \langle -Ax^{(0)}, -A^2 x^{(0)} \rangle$$

$$\|r^{(0)}\|_A^2 = \langle r^{(0)}, A r^{(0)} \rangle = \langle -Ax^{(0)}, -A^2 x^{(0)} \rangle$$
$$= (Ax^{(0)})^T A^2 x^{(0)} = x^{(0)T} A^T A^2 x^{(0)}$$
$$\|r^{(0)}\|_A^2 = x^{(0)T} A^3 x^{(0)}$$

$$\therefore \quad \frac{\|r^{(0)}\|_2^2}{\|r^{(0)}\|_A^2} = \frac{x^{(0)T} A^2 x^{(0)}}{x^{(0)T} A^3 x^{(0)}}$$
$$= \frac{a^2(1^2) + 1^2(a^2)}{a^2(1^3) + 1^2(a^3)} = \frac{a^2(1+1)}{a^2(1+a)}$$
$$\alpha^{(0)} = \frac{2}{1+a}$$

So,
$$\alpha^{(0)} r^{(0)} = \frac{-2}{1+a} A x^{(0)}$$

$$\therefore \quad x^{(1)} = x^{(0)} - \frac{2}{1+a} A x^{(0)} = \left(I - \frac{2}{1+a} A\right) x^{(0)}$$

$$I - \frac{2}{1+a} A = \mathrm{diag}\left(1 - \frac{2}{1+a}, 1 - \frac{2a}{1+a}\right)$$
$$= \mathrm{diag}\left(\frac{a-1}{a+1}, \frac{1-a}{a+1}\right)$$
$$\therefore \quad x^{(1)} = \frac{a-1}{a+1} \mathrm{diag}(1,-1)\, x^{(0)}, \text{ let } \mu = \frac{a-1}{a+1}$$

$$x^{(2)} = x^{(1)} + \alpha^{(1)} r^{(1)}$$

$$r^{(1)} = b - Ax^{(1)} = -Ax^{(1)}$$
$$\|r^{(1)}\|_2^2 = x^{(1)T} A^2 x^{(1)}$$

$$\|r^{(1)}\|_A^2 = x^{(1)T} A^3 x^{(1)}$$

$$\left\{ \begin{aligned} &\text{In general } \|r^{(k)}\|_2^2 = x^{(k)T} A^2 x^{(k)}, \\ &\|r^{(k)}\|_A^2 = x^{(k)T} A^3 x^{(k)} \end{aligned} \right.$$

$$\alpha^{(1)} = \frac{\mu \; x^{(0)T} \operatorname{diag}(1,-1) A^2 \mu \operatorname{diag}(1,-1) x^{(0)}}{\mu \; x^{(0)T} \operatorname{diag}(1,-1) A^3 \mu \operatorname{diag}(1,-1) x^{(0)}}$$

Note that $\operatorname{diag}(1,-1) B \operatorname{diag}(1,-1)$ only negates the anti-diagonal entries. Also, μ is scalar.

$$\alpha^{(1)} = \frac{\mu^2 x^{(0)T} A^2 x^{(0)}}{\mu^2 x^{(0)T} A^3 x^{(0)}} = \frac{a^2(1^2) + 1^2(a^2)}{a^2(1^3) + 1^2(a^3)}$$

$$\alpha^{(1)} = \frac{2}{1+a}$$

$$x^{(2)} = x^{(1)} - \frac{2}{1+a} A x^{(1)}$$

$$x^{(2)} = \left(I - \frac{2}{1+a} A\right) x^{(1)}$$

$$x^{(2)} = \mu \operatorname{diag}(1,-1) \mu \operatorname{diag}(1,-1) x^{(0)}$$

$$x^{(2)} = \mu^2 x^{(0)}$$

Note that $\operatorname{diag}(1,-1)\operatorname{diag}(1,-1) = I_2$

$$x^{(3)} = x^{(2)} + \alpha^{(2)} r^{(2)}$$

$$\alpha^{(2)} = \frac{x^{(2)T} A^2 x^{(2)}}{x^{(2)T} A^3 x^{(2)}} = \frac{\mu^2 x^{(0)T} A^2 \mu^2 x^{(0)}}{\mu^2 x^{(0)T} A^3 \mu^2 x^{(0)}}$$

$$= \frac{x^{(0)T} A^2 x^{(0)}}{x^{(0)T} A^3 x^{(0)}} = \frac{2}{1+a}$$

$$x^{(3)} = x^{(2)} - \frac{2}{1+a} A x^{(2)}$$

$$x^{(3)} = \left(I - \frac{2}{1+a} A \right) x^{(2)}$$

$$x^{(3)} = \mu \, \mathrm{diag}(1,-1) \mu^2 x^{(0)}$$

$$x^{(3)} = \mu^3 \, \mathrm{diag}(1,-1) x^{(0)}$$

We can notice a pattern here, that is

$$\alpha^{(k)} = \frac{\mu^k x^{(0)T} A^2 \mu^k x^{(0)}}{\mu^k x^{(0)T} A^3 \mu^k x^{(0)}} = \frac{2}{1+a}, \forall k.$$

$$x^{(k)} = \left(I - \frac{2}{1+a} A \right) x^{(k-1)}$$

$$x^{(k)} = \mu \, \mathrm{diag}(1,-1) x^{(k-1)}$$

$$x^{(k)} = \mu^k \, \mathrm{diag}(1,-1)^k x^{(0)}$$

When k is even ie $k = 2m$

$$x^{(2m)} = \mu^{2m} x^{(0)} = \left(\frac{a-1}{a+1} \right)^{2m} x^{(0)},$$

when k is odd i.e $k = 2m+1$,

$$x^{(2m+1)} = \mu^{2m+1} \operatorname{diag}(1,-1) x^{(0)}$$
$$= \left(\frac{a-1}{a+1}\right)^{2m+1} \operatorname{diag}(1,-1) x^{(0)}$$

$$r^{(k)} = -Ax^{(k)}$$
$$= -A\mu^{k} \operatorname{diag}(1,-1)^{k} x^{(0)}$$

$$r^{(k)} = \begin{cases} -\mu^{k} A \begin{bmatrix} a \\ 1 \end{bmatrix}, & \text{if } k \text{ is even} \\ -\mu^{k} A \begin{bmatrix} a \\ -1 \end{bmatrix}, & \text{if } k \text{ is odd} \end{cases}$$

(b) $$\cos(\phi) = \frac{\langle x, y \rangle_A}{\|x\|_A \|y\|_A}$$

$$\langle r^{(k)}, r^{(k+1)} \rangle_A = r^{(k)T} A r^{(k+1)}$$

$$= \mu^{2k+1} [a \quad 1] A^T A^2 \begin{bmatrix} a \\ -1 \end{bmatrix}$$

$$= \mu^{2k+1} [a \quad 1] A^3 \begin{bmatrix} a \\ -1 \end{bmatrix}$$

$$= \mu^{2k+1} [a^2(1^3) + (1\times -1)(a^3)]$$
$$= \mu^{2k+1} (a^2 - a^3)$$

$$\|r^{(k)}\|_A^2 = r^{(k)T} A r^{(k)} = \mu^{2k} [a \quad 1] A^T A A \begin{bmatrix} a \\ 1 \end{bmatrix}$$

$$\|r^{(k)}\|_A^2 = \mu^{2k} [a \quad 1] A^3 \begin{bmatrix} a \\ 1 \end{bmatrix}$$

$$= \mu^{2k}(a^2 (1^3) + 1^2 (a^3))$$

$$= \mu^{2k} (a^2 + a^3)$$

$$\|r^{(k+1)}\|_A^2 = \mu^{2k+2} [a \quad -1] A^T A A \begin{bmatrix} a \\ -1 \end{bmatrix}$$

$$= \mu^{2k+2} [a \quad -1] A^3 \begin{bmatrix} a \\ -1 \end{bmatrix}$$

$$= \mu^{2k+2} [a^2 (1^3) + (-1)^2 (a^3)]$$

$$= \mu^{2k+2} (a^2 + a^3)$$

$\Rightarrow$

$$\|r^{(k)}\|_A = \mu^k \sqrt{a^2 + a^3}$$

$$\|r^{(k+1)}\|_A = \mu^{k+1} \sqrt{a^2 + a^3}$$

$$\therefore \quad \frac{\langle r^{(k)}, r^{(k+1)} \rangle_A}{\|r^{(k)}\|_A \|r^{(k+1)}\|_A} = \frac{\mu^{2k+1} (a^2 - a^3)}{\mu^{2k+1} (a^2 + a^3)}$$

$$\cos(\phi) = \frac{a^2 - a^3}{a^2 + a^3} = \frac{a^2 (1 - a)}{a^2 (1 + a)}$$

$$\cos(\phi) = \frac{1-a}{1+a} = \frac{1/a - 1}{1/a + 1}$$ ← multiply both numerator and denominator by $1/a$

For large enough a i.e $a \to \infty$

$$\cos(\phi) = \frac{0-1}{0+1} = -1$$

$$\boxed{\phi = \cos^{-1}(-1) = \pi}$$