TMA4205 Numerical Linear Algebra Fall 2013

Norwegian University of Science and Technology
Department of Mathematical Sciences

Solutions to exercise set 4

1 **Saad, Exercise 5.3** We first consider the system $Ax = b$ with $\mathscr{L} = A\mathscr{K}$. Let

$$\mathscr{K} = \operatorname{span}\{v_1, \dots, v_m\},$$
$$V_m = [v_1 \mid \cdots \mid v_m] \in \mathbb{R}^{n\times m}.$$

In the first case, the approximate solution $\tilde{x}$ to $Ax = b$ must satisfy

$$\tilde{x} - x_0 \in \mathscr{K},$$
$$(AV_m)^{\mathrm{T}}(b - A\tilde{x}) = 0.$$

Next, we consider the system $A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$ with $\mathscr{L} = \mathscr{K}$ (*orthogonal* projection). The approximate solution must now satisfy

$$\tilde{x} - x_0 \in \mathscr{K},$$
$$V_m^{\mathrm{T}}(A^{\mathrm{T}}b - A^{\mathrm{T}}A\tilde{x}) = 0$$
$$\Downarrow$$
$$(AV_m)^{\mathrm{T}}(b - A\tilde{x}) = 0.$$

Hence, we get the same system $(AV_m)^{\mathrm{T}}A\tilde{x} = (AV_m)^{\mathrm{T}}b$ or $(AV_m)^{\mathrm{T}}A\delta = (AV_m)^{\mathrm{T}}r_0$, where $\tilde{x} = x_0 + \delta$. This shows that the two methods are equivalent.

2 We use $N = 100$, $v = \mathrm{e}_1$, and $m = 10, 20, 30, 40, 50$, as suggested. The results for both **a)** and **b)** are reported in Table 1, where we report the errors $\|V_m^{\mathrm{T}}AV_m - H_m\|_\infty$ and $\|V_m^{\mathrm{T}}V_m - \mathrm{I}_m\|_\infty$ after $m = 10, 20, 30, 40, 50$ iterations for Gram–Schmidt (GS) orthogonalization and modified Gram–Schmidt (MGS) orthogonalization. From the table we conclude that MGS performs slightly better than GS in this case.

	GS		MGS	
m	$\|V_m^{\mathrm{T}}AV_m - H_m\|_\infty$	$\|V_m^{\mathrm{T}}V_m - \mathrm{I}_m\|_\infty$	$\|V_m^{\mathrm{T}}AV_m - H_m\|_\infty$	$\|V_m^{\mathrm{T}}V_m - \mathrm{I}_m\|_\infty$
10	$2.90 \cdot 10^{-13}$	$3.91 \cdot 10^{-14}$	$9.69 \cdot 10^{-14}$	$1.92 \cdot 10^{-14}$
20	$1.33 \cdot 10^{-11}$	$1.72 \cdot 10^{-12}$	$3.89 \cdot 10^{-12}$	$7.73 \cdot 10^{-13}$
30	$2.04 \cdot 10^{-8}$	$2.64 \cdot 10^{-9}$	$5.98 \cdot 10^{-9}$	$1.19 \cdot 10^{-9}$
40	$3.55 \cdot 10^{-3}$	$4.61 \cdot 10^{-4}$	$1.04 \cdot 10^{-3}$	$2.07 \cdot 10^{-4}$
50	80.4	8.04	19.2	2.41

Table 1: MGS is a little bit better than GS.

3 **a)** Since A is SPD, it follows from Exercise 2, Problem **1c)**, that A^{-1} is also SPD. Thus, it can be used to define the A^{-1}-norm in the same way as the SPD matrix A defines the A-norm.

b) We are given the system $Ax = b$, where A is SPD. Then, for all $\tilde{x} \in \mathbb{R}^n$,

$$\begin{aligned}
\|x-\tilde{x}\|_A^2 &= (x-\tilde{x})^{\mathrm{T}} A (x-\tilde{x}) \\
&= (x-\tilde{x})^{\mathrm{T}} A A^{-1} A (x-\tilde{x}) \\
&= (Ax - A\tilde{x})^{\mathrm{T}} A^{-1} (Ax - A\tilde{x}) \\
&= (b - A\tilde{x})^{\mathrm{T}} A^{-1} (b - A\tilde{x}) \\
&= \tilde{r}^{\mathrm{T}} A^{-1} \tilde{r} \\
&= \|\tilde{r}\|_{A^{-1}}^2 .
\end{aligned}$$

Hence,

$$x_m = \underset{\tilde{x}\in\mathcal{K}_m}{\operatorname{argmin}} \|x-\tilde{x}\|_A^2 = \underset{\tilde{x}\in\mathcal{K}_m}{\operatorname{argmin}} \|\tilde{r}\|_{A^{-1}}^2 .$$

c) The functional $f\colon \mathbb{R}^n \to \mathbb{R}$ is given by $f(w) = \frac{1}{2} w^{\mathrm{T}} A w - w^{\mathrm{T}} b$. This functional attains a minimum since A is SPD. Starting at x_j and minimizing f along the search direction p_j gives

$$\begin{aligned}
f(x_j + \alpha p_j) &= \tfrac{1}{2}(x_j + \alpha p_j)^{\mathrm{T}} A (x_j + \alpha p_j) - (x_j + \alpha p_j)^{\mathrm{T}} b, \\
&= \tfrac{1}{2} x_j^{\mathrm{T}} A x_j + \alpha x_j^{\mathrm{T}} A p_j + \tfrac{1}{2}\alpha^2 p_j^{\mathrm{T}} A p_j - x_j^{\mathrm{T}} b - \alpha p_j^{\mathrm{T}} b.
\end{aligned}$$

At the minimum,

$$\begin{aligned}
0 = \frac{\mathrm{d}f}{\mathrm{d}\alpha} &= x_j^{\mathrm{T}} A p_j + \alpha p_j^{\mathrm{T}} A p_j - p_j^{\mathrm{T}} b \\
&= -p_j^{\mathrm{T}} r_j + \alpha p_j^{\mathrm{T}} A p_j \\
&\Downarrow \\
\alpha &= \frac{(p_j, r_j)}{(Ap_j, p_j)}.
\end{aligned}$$

But

$$\begin{aligned}
p_j &= r_j + \beta_{j-1} p_{j-1} \\
&= r_j + \beta_{j-1}(r_{j-1} + \beta_{j-2} p_{j-2}) \\
&\vdots \\
&= r_j + \sum_{i=0}^{j-1} c_i r_i ,
\end{aligned}$$

where c_i are constants. Hence, due to the orthogonality of the residuals r_i, we get that $(p_j, r_j) = (r_j, r_j)$, which in turn gives us the result

$$\alpha = \frac{(r_j, r_j)}{(Ap_j, p_j)}.$$