1) In class, we showed that

$$p_{k+1} = r_{k+1} - \frac{\langle p_k, r_{k+1}\rangle_A}{\|p_k\|_A^2} p_k. \tag{1}$$

(a) Using the fact that $r_{k+1} = r_k - \alpha_k A p_k$ and $r_{k+1}^T r_k = 0$, show that $\langle p_k, r_{k+1}\rangle_A = -\frac{\|r_{k+1}\|_2^2}{\alpha_k}$.

$$\begin{aligned}
0 = r_{k+1}^T r_k &= r_{k+1}^T (r_{k+1} + \alpha_k A p_k) \\
&= r_{k+1}^T r_{k+1} + \alpha_k r_{k+1} A p_k \\
\implies r_{k+1} A p_k &= -\frac{r_{k+1}^T r_{k+1}}{\alpha_k}
\end{aligned}$$

which implies

$$\langle p_k, r_{k+1}\rangle_A = -\frac{\|r_{k+1}\|_2^2}{\alpha_k}.$$

(b) Rewrite $\|p_k\|_A^2$ in terms of r_k and α_k.

$$\begin{aligned}
\|p_k\|_A^2 &= p_k^T A p_k \\
&= \left(r_k - \frac{\langle p_{k-1}, r_k\rangle}{\|p_{k-1}\|_A^2} p_{k-1} \right)^T \frac{1}{\alpha_k}(r_k - r_{k+1}) \\
&= \frac{1}{\alpha_k}(r_k^T r_k - r_k^T r_{k+1}) \quad \text{because } p_{k-1} \text{ is orthogonal to } r_k \text{ and } r_{k+1} \\
&= \frac{1}{\alpha_k} r_k^T r_k \\
&= \frac{1}{\alpha_k} \|r_k\|_2^2.
\end{aligned}$$

(c) Plug these expressions into (1) to get a technique for evaluating the next basis vector for the residual space without any applications of the matrix A.

$$\begin{aligned}
p_{k+1} &= r_{k+1} - \left(-\frac{\|r_{k+1}\|_2^2}{\alpha_k} \right) \left(\frac{\alpha_k}{\|r_k\|_2^2} \right) p_k \\
&= r_{k+1} + \left(\frac{\|r_{k+1}\|_2}{\|r_k\|_2} \right)^2 p_k.
\end{aligned}$$

2) Consider a sparse 500×500 matrix A constructed as follows.

- Put a 1 in each diagonal entry.
- In each off-diagonal entry put a random number from the uniform distribution on $[-1, 1]$ but make sure to maintain symmetry. Then replace each off-diagonal entry with $|A_{ij}| > \tau$ by 0, where τ for $\tau = 0.01, 0.05, 0.1$, and 0.2.

Take the right hand side to be a random vector b and set the tolerance to 10^{-10}.

(a) Write the Steepest Descent (SD) and Conjugate Gradient (CG) solver.

(b) Apply SD to solve each of the linear systems ad plot the residual for each iteration $\|r_n\|$ versus the iteration n on a *semilogy* scale.

(c) Apply CG to solve each of the linear systems ad plot the residual for each iteration $\|r_n\|$ versus the iteration n on a *semilogy* scale.

(d) What do you observe about the convergence of these methods? If the methods do not converge for any choices of τ explain what's happening.

(e) How do the residual compare with the error bounds provided in class?

3) Suppose CG is applied to a symmetric positive definite matrix A with the result $\|e_0\|_A = 1$, and $\|e_{10}\|_A = 2 \cdot 2^{-10}$, where $\|e_k\|_A = \|x_k - x^*\|_A$ and x^* is the true solution. Based solely on this data,

(a) What bound can you give on $\kappa(A)$?

(b) What bound can you give on $\|e_{20}\|_A$?

4) Consider the task of solving the following system of nonlinear equations.

$$f_1(x, y) = 3x^2 + 4y^2 - 1 = 0 \text{ and } f_2(x, y) = y^3 - 8x^3 - 1 = 0$$

for the solution α near $(x, y) = (-0.5, 0.25)$.

(a) Apply the fixed point iteration with

$$g(x) = x - \begin{pmatrix} 0.016 & -0.17 \\ 0.52 & -0.26 \end{pmatrix} \begin{pmatrix} 3x^2 + 4y^2 - 1 \\ y^3 - 8x^3 - 1 \end{pmatrix}.$$

You can use $(-0.5, 0.25)$ as the initial condition. How many steps are needed to get an approximation to 7 digits of accuracy?

(b) Why is this a good choice for $g(x)$.