Math 151AH HW 3 Andrea Barretta

1) a. $x^{(k+1)} = x^{(k)} + \alpha(b - Ax^{(k)})$

$= \alpha(\frac{1}{\alpha} I x^{(k)}) + \alpha(b - Ax^{(k)})$

$= \alpha(b - (A - \frac{1}{\alpha} I)x^{(k)})$

Then $A_2^{-1} = \alpha I$, so $A_2 = \frac{1}{\alpha} I$

$\Rightarrow A_1 = A - \frac{1}{\alpha} I$

b. Let $x^{(k+1)} = G(x^{(k)}) = x^{(k)} + \alpha(b - Ax^{(k)})$

We need G to be a contraction mapping.

$\|G(x) - G(y)\|_2 = \| x + \alpha(b - Ax) - y - \alpha(b - Ay)\|_2$

$= \| x - y + \alpha A(y - x)\|_2$

$= \| x - y - \alpha A(x - y)\|_2$

$= \| (I - \alpha A)(x - y)\|_2$

$\leq \|I - \alpha A\|_2 \|x - y\|_2$

Let (λ, x) be an eigenvalue - eigenvector pair of A. Then

$(I - \alpha A)(x) = x - \alpha\lambda x = (1 - \alpha\lambda)x$

So $1 - \alpha\lambda$ is an eigenvalue of $I - \alpha A$. Since A is symmetric, so is $I - \alpha A$. Then $\|I - \alpha A\|_2 = \rho(I - \alpha A)$, which will be either $|1 - \alpha\lambda_1|$ or $|1 - \alpha\lambda_n|$. Then

$|1 - \alpha\lambda_1| < 1$ $\qquad$ $|1 - \alpha\lambda_n| < 1$

$-2 < -\alpha\lambda_1 < 0$ $\qquad$ $-2 < -\alpha\lambda_n < 0$

$\frac{2}{\lambda_1} > \alpha > 0$ $\qquad$ $\frac{2}{\lambda_2} > \alpha > 0$

Then $\alpha \in (0, \frac{2}{\lambda_1}) \cap (0, \frac{2}{\lambda_2}) = (0, \frac{2}{\lambda_1})$ since $\frac{2}{\lambda_1} < \frac{2}{\lambda_2}$. Then we need $\alpha \in (0, \frac{2}{\lambda_1})$

c. From (c), $x^{(k+1)} = \alpha(b - (A - \frac{1}{\alpha} I)x^{(k)})$

$= \alpha b + [-(\alpha A - I)]x^{(k)}$

$= \alpha b + \underbrace{[I - \alpha A]}_{\text{iteration matrix}} x^{(k)}$

From earlier, the eigenvalues of $I - \alpha A$ are $1 - \alpha\lambda_1, \ldots, 1 - \alpha\lambda_n$. Then we seek to minimize $\rho(I - \alpha A) = \max\{|1 - \alpha\lambda_1|, \ldots, |1 - \alpha\lambda_n|\} = \max\{|1 - \alpha\lambda_1|, |1 - \alpha\lambda_n|\}$. We can do so by centering the two at 0, such that $1 - \alpha\lambda_1 < 0$, so we want to impose that

$$|1-\alpha\lambda_1| = |1-\alpha\lambda_n|$$
$$\lambda_1\alpha - 1 = 1-\alpha\lambda_n$$
$$(\lambda_1+\lambda_2)\alpha = 2$$
$$\alpha = \frac{2}{\lambda_1+\lambda_2}$$

As desired.

d. Let (λ, x) be an eigenvalue-eigenvector pair of A. Then

$$\lambda x = Ax$$
$$\lambda A^{-1}x = x \Rightarrow A^{-1}x = \frac{1}{\lambda}x$$

Then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}, so the eigenvalues of A^{-1} are $\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$. Further, since A is symmetric, it is orthogonally diagonalizable, so there exists a unitary W, diagonal D such that

$$A = WDW^*$$
$$A^{-1} = W^{*^{-1}}D^{-1}W^{-1} = WD^{-1}W^*$$

Then A^{-1} is a real, orthogonally diagonalizable matrix, so A is symmetric. Then $\|A^{-1}\|_2 = \rho(A^{-1}) = \frac{1}{\lambda_n}$.

Then $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 = \lambda_1\frac{1}{\lambda_n}$. Further,

$$M = -A_2^{-1}A_1$$
$$= -(\alpha I)(A - \tfrac{1}{\alpha}I)$$
$$= -(\alpha A - I) = I - \alpha A$$

From (c), for $\alpha = \frac{2}{\lambda_1+\lambda_n}$,

$$\rho(M) = \rho(I-\alpha A) = \alpha\lambda_1 - 1 = \frac{2\lambda_1}{\lambda_1+\lambda_n} - 1$$
$$= \frac{2\kappa_2(A)\lambda_n}{\kappa_2(A)\lambda_n+\lambda_n} - 1$$
$$= \frac{2\kappa_2(A)}{\kappa_2(A)+1} - \frac{\kappa_2(A)+1}{\kappa_2(A)+1}$$
$$= \frac{\kappa_2(A)-1}{\kappa_2(A)+1}$$

2) a. $G(x_1, x_2) = \begin{bmatrix} \gamma(x_2 - x_1)\cos(\frac{\pi}{10}) - \gamma(x_2 - 1)\sin(\frac{\pi}{10}) + 1 \\ \gamma(x_1 - 1)\sin(\frac{\pi}{10}) + \gamma(x_2 - 1)\cos(\frac{\pi}{10}) + 5(x_1 - 1)^3 + 1 \end{bmatrix}$

$G(x^*) = G(1,1) = \begin{bmatrix} 0 - 0 + 1 \\ 0 + 0 + 0 + 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = x^*$

b. $\frac{\partial G_1}{\partial x_1} = \gamma\cos(\frac{\pi}{10}) \qquad \frac{\partial G_1}{\partial x_2} = -\gamma\sin(\frac{\pi}{10})$

$\frac{\partial G_2}{\partial x_1} = \gamma\sin(\frac{\pi}{10}) + 15(x-1)^2 \qquad \frac{\partial G_2}{\partial x_2} = \gamma\cos(\frac{\pi}{10})$

Then

$\nabla G(x^*) = \begin{bmatrix} \gamma\cos(\frac{\pi}{10}) & -\gamma\sin(\frac{\pi}{10}) \\ \gamma\sin(\frac{\pi}{10}) & \gamma\cos(\frac{\pi}{10}) \end{bmatrix}$

c. $\|\nabla G(x^*)\|_2 = |\lambda| \Bigg\| \underbrace{\begin{bmatrix} \cos\frac{\pi}{10} & -\sin\frac{\pi}{10} \\ \sin\frac{\pi}{10} & \cos\frac{\pi}{10} \end{bmatrix}}_{\text{Rotation matrix, orthogonal}} \Bigg\|_2$

$= |\lambda| \sup_{\substack{x \in \mathbb{R}^2 \\ x \neq 0}} \frac{\left\| \begin{bmatrix} \cos\frac{\pi}{10} & -\sin\frac{\pi}{10} \\ \sin\frac{\pi}{10} & \cos\frac{\pi}{10} \end{bmatrix} x \right\|}{\|x\|} = |\lambda| \sup_{\substack{x \in \mathbb{R}^2 \\ x \neq 0}} \frac{\|x\|}{\|x\|} = |\lambda|$

d. Then fixed point iteration converges for $|\lambda| < 1$.

3) a. $F(x_1, x_2) = \begin{bmatrix} x_1^2 + x_2^2 + 5x_1 \\ 2x_1x_2 + 3x_2^2 + x_2 \end{bmatrix}$

$F(x^*) = F(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = x^*$

$\frac{\partial F_1}{\partial x_1} = 2x_1 + 5 \qquad \frac{\partial F_1}{\partial x_2} = 2x_2$

$\frac{\partial F_2}{\partial x_1} = 2x_2 \qquad \frac{\partial F_2}{\partial x_2} = 2x_1 + 6x_2 + 1$

Then

$$\nabla F(x_1, x_2) = \begin{bmatrix} 2x_1 + 5 & 2x_2 \\ 2x_2 & 2x_1 + 6x_2 + 1 \end{bmatrix}$$

b. $\|\nabla G(x^*)\|_2 = \| I - \alpha \nabla F(x^*) \|_2$

$= \left\| I - \alpha \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix} \right\|_2$

$= \left\| \begin{bmatrix} 1-\alpha 5 & 0 \\ 0 & 1-\alpha \end{bmatrix} \right\|_2 = \rho\left(\begin{bmatrix} 1-\alpha 5 & 0 \\ 0 & 1-\alpha \end{bmatrix}\right)$

By symmetry. Then we need

$|1-5\alpha| < 1 \qquad |1-\alpha| < 1$

$-2 < -5\alpha < 0 \qquad -2 < -\alpha < 0$

$\frac{2}{5} > \alpha > 0 \qquad 2 > \alpha > 0$

So we need $0 < \alpha < \frac{2}{5}$

4) Sherman - Morrison - $(A + xy^T)^{-1} = A^{-1} - \dfrac{A^{-1}xy^TA^{-1}}{1+y^TA^{-1}x}$

From Broyden's,

$$A_k = A_{k-1} + \frac{1}{\|s^{(k)}\|_2^2}\left[y^{(k)} - A_{k-1}s^{(k)}\right]s^{(k)T}$$

$$A_k^{-1} = \left(A_{k-1} + \frac{1}{\|s^{(k)}\|_2^2}\left[y^{(k)} - A_{k-1}s^{(k)}\right]s^{(k)T}\right)^{-1}$$

$$= A_{k-1}^{-1} - \frac{A_{k-1}^{-1}\frac{1}{\|s^{(k)}\|_2^2}(y^{(k)} - A_{k-1}s^{(k)})s^{(k)T}A_{k-1}^{-1}}{1 + \frac{1}{\|s^{(k)}\|_2^2}(y^{(k)} - A_{k-1}s^{(k)})^TA_{k-1}^{-1}s^{(k)}}$$

$$= A_{k-1}^{-1} - \frac{A_{k-1}^{-1}(y^{(k)} - A_{k-1}s^{(k)})s^{(k)T}A_{k-1}^{-1}}{\|s^{(k)}\|_2^2 + (y^{(k)T} - s^{(k)T}A_{k-1}^T)A_{k-1}^{-1}s^{(k)}}$$

$$= A_{k-1}^{-1} - \frac{A_{k-1}^{-1}y^{(k)}s^{(k)T}A_{k-1}^{-1} - s^{(k)}s^{(k)T}A_{k-1}^{-1}}{\|s^{(k)}\|_2^2 + y^{(k)T}A_{k-1}^{-1}s^{(k)} - \|s^{(k)}\|_2^2}$$

$$= A_{k-1}^{-1} + \left(\frac{s^{(k)} - A_{k-1}^{-1}y^{(k)}}{y^{(k)T}A_{k-1}^{-1}s^{(k)}}\right)s^{(k)T}A_{k-1}^{-1}$$