Sample Solutions for Practice Problems for Midterm (Fri., Feb. 29).

1. Let N be the random variable which counts the number of books a library patron checks out. The library limits the maximum number of books one can check out to 3. Suppose the probability distribution of N is determined by

1	2	3
$\frac{2}{10}$	$\frac{5}{10}$	$\frac{3}{10}$

(a) How many books on average does a patron take out?

$$E(N) = \frac{2}{10} \cdot 1 + \frac{5}{10} \cdot 2 + \frac{3}{10} \cdot 3 = \frac{21}{10}.$$

(b) The library has data on the number of patrons visiting each day and the total number of books borrowed each day for the past year. About 1000 patrons visit the library each day. What should the histogram of the daily average number of books per customer look like? Sketch it, labeling important points, and explain your reasoning.

The histogram should be *normally* distributed, with mean 21/10 and variance $\sigma^2/1000$, where

$$\sigma^2 = E(N^2) - (E(N))^2 = \frac{2}{10} \cdot 1^2 + \frac{5}{10} \cdot 2^2 + \frac{3}{10} \cdot 3^2 - \left(\frac{21}{10}\right)^2 = \frac{49}{10} - \frac{441}{100} = \frac{49}{100}.$$

This follows from the Central Limit Theorem. The standard deviation of the daily averages will be about $\sqrt{49/100000} \approx .022$, and about 65% of the daily averages will lie between $2.1 - .022$ and $2.1 + .022$.

2. Suppose we have a uniform random number generator, such as `rand` in MATLAB, that generates random numbers between 0 and 1 from a uniform distribution. We wish to generate random numbers from a distribution X, where

$$\text{prob}(X \leq a) = \sqrt{a}, \quad 0 \leq a \leq 1.$$

How could we use the uniform random number generator to obtain random numbers with this distribution?

For a uniform distribution, $\text{prob}(Y \leq a) = a$, for $0 \leq a \leq 1$. If $X = Y^2$, then $\text{prob}(X \leq a) = \text{prob}(Y \leq \sqrt{a}) = \sqrt{a}$. Therefore, set $X = Y^2$, where Y comes from the uniform random number generator.

3. Suppose a real symmetric matrix A has eigenvalues -4, -2, 1, 3, and 5. Assume that the initial vector in the following algorithms has nonzero components in the direction of each eigenvector.

(a) To which eigenvalue (if any) will the unshifted power method converge?

The one with largest absolute value, 5.

(b) Derive an expression showing the *rate of convergence* of the power method.

Let the eigenvalues be labelled $\lambda_1 \leq \ldots \leq \lambda_5$, and let $\mathbf{v_1}, \ldots, \mathbf{v}_5$ be the corresponding eigenvectors. Write the initial vector $\mathbf{w}$ as $\mathbf{w} = \sum_{j=1}^{5} c_j \mathbf{v_j}$, for certain coefficients c_j. Then

$$A^k \mathbf{w} = \sum_{j=1}^{5} c_j \lambda_j^k \mathbf{v_j},$$

and, dividing each side by $\lambda_5^k = 5^k$,

$$\frac{1}{5^k} A^k \mathbf{w} = c_5 \mathbf{v}_5 + \sum_{j=1}^{4} c_j \left(\frac{\lambda_j}{5} \right)^k \mathbf{v}_j.$$

Since each fraction $\lambda_j/5$, $j = 1, \ldots, 4$, is less than 1 in absolute value, it follows that $\frac{1}{5^k} A^k \mathbf{w}$ converges to $c_5 \mathbf{v_5}$ as $k \to \infty$. The *rate of convergence* is determined by the largest ratio $|\lambda_j/5|$, $j = 1, \ldots, 4$, which is $|\lambda_1/5| = 4/5$.

(c) To which eigenvalue (if any) will unshifted inverse iteration converge?

The one closest to the origin, 1.

(d) Suppose you wish to compute the eigenvector corresponding to the eigenvalue 3. What range of shifts could you use with inverse iteration to make it converge to that eigenvector?

Need a shift s for which $|3-s| < \min\{|5-s|, |1-s|\}$. Need $s < 4$ in order to have $|3-s| < |5-s|$, and need $s > 2$ in order to have $|3-s| < |1-s|$. Therefore $s \in (2, 4)$.

4. Consider the following matrix:

$$A = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

(a) Compute the QR factorization of A. Explicitly write down the orthogonal matrix Q and the upper triangular matrix R, and show how you found them.

Normalize the first column of A to get

$$\mathbf{q_1} = \begin{pmatrix} 4/5 \\ -3/5 \end{pmatrix}.$$

Orthogonalize the second column of A against $\mathbf{q_1}$ to get

$$\tilde{\mathbf{q}}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \frac{7}{5}\begin{pmatrix} 4/5 \\ -3/5 \end{pmatrix} = \begin{pmatrix} 3/25 \\ 4/25 \end{pmatrix},$$

and normalize this to get

$$\mathbf{q_2} = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}.$$

Let $Q = (\mathbf{q}_1, \mathbf{q_2})$. Then $A = QR$, where

$$Q = \begin{pmatrix} 4/5 & 3/5 \\ -3/5 & 4/5 \end{pmatrix}, \quad R = \begin{pmatrix} 5 & -7/5 \\ 0 & 1/5 \end{pmatrix}.$$

(b) Write down the first step of the QR algorithm for A; that is, compute the first matrix A_1 that is similar to A in the QR algorithm.

$$A_1 = RQ = \begin{pmatrix} 5 & -7/5 \\ 0 & 1/5 \end{pmatrix}\begin{pmatrix} 4/5 & 3/5 \\ -3/5 & 4/5 \end{pmatrix} = \frac{1}{25}\begin{pmatrix} 121 & 47 \\ -3 & 4 \end{pmatrix}.$$

(c) Explain what information Gerschgorin's theorem provides about the eigenvalues of A. Draw a graph showing the region(s) in which the eigenvalues lie.

The Gerschgorin row disks tell us that the two eigenvalues of A lie in the union of: the disk of radius 1 about 4 and the disk of radius 3 about 1. Since these disks overlap, we do not know anything about where in this union of disks the two eigenvalues lie – they could both be in the same disk or they could be in different disks.

The Gerschgorin column disks tell us that the two eigenvalues of A lie in the union of: the disk of radius 3 about 4 and the disk of radius 1 about 1. Again the disks overlap, so we cannot say anything about which disk contains which eigenvalue or if one disk contains both of them.

Combining the information from the row and column disks, if $D(r, c)$ denotes the disk of radius r about c, then we know that the eigenvalues of A lie in $(D(1,4) \cup D(3,1)) \cap (D(3,4) \cup D(1,1))$.

5. Consider the shifted QR algorithm. Suppose $A_k - s_k I = Q_k R_k$ and $A_{k+1} = R_k Q_k + s_k I$. Show that A_{k+1} is orthogonally similar to A_k. Be sure to justify all steps.

$$\begin{aligned}
A_{k+1} &= (Q_k^T Q_k) R_k Q_k + s_k I && \text{since } Q_k^T Q_k = I \\
&= Q_k^T (Q_k R_k) Q_k + s_k I && \text{since matrix mult is associative} \\
&= Q_k^T (A_k - s_k I) Q_k + s_k I && \text{since } A_k - s_k I = Q_k R_k \\
&= Q_k^T A_k Q_k - s_k Q_k^T I Q_k + s_k I && \text{since matrix mult is distributive} \\
&= Q_k^T A_k Q_k - s_k I + s_k I && \text{since } Q_k^T I Q_k = I \\
&= Q_k^T A_k Q_k.
\end{aligned}$$

Thus A_{k+1} is similar to A_k via the orthogonal matrix Q_k.

6. Consider a derivative approximation of the form

$$f'(x) \approx Af(x) + Bf(x-h) + Cf(x-2h).$$

(a) Use Taylor's theorem to determine coefficients A, B, and C for which this approximation is second order accurate. Justify your answer.

Expand $f(x-h)$ and $f(x-2h)$ in Taylor series about x:

$$\begin{aligned} f(x-h) &= f(x) - hf'(x) + \frac{h^2}{2}f''(x) + O(h^3) \\ f(x-2h) &= f(x) - 2hf'(x) + \frac{(2h)^2}{2}f''(x) + O(h^3) \end{aligned}$$

Add $Af(x) + Bf(x-h) + Cf(x-2h)$ to find:

$$Af(x)+Bf(x-h)+Cf(x-2h) = (A+B+C)f(x)-(B+2C)hf'(x)+(B+4C)\frac{h^2}{2}f''(x)$$

$$+O((|B|+|C|)h^3).$$

In order for this to approximate $f'(x)$, we need:

$$\begin{aligned} A+B+C &= 0 \\ -(B+2C)h &= 1 \\ B+4C &= 0. \end{aligned}$$

The solution is

$$C = \frac{1}{2h}, \quad B = -\frac{2}{h}, \quad A = \frac{3}{2h},$$

and then we have

$$\frac{3}{2h}f(x) - \frac{2}{h}f(x-h) + \frac{1}{2h}f(x-2h) = f'(x) + O(h^2).$$

(b) Suppose you use your formula from (a) to generate an approximation $A_{.01}$ with $h = .01$ and another approximation $A_{.005}$ with $h = .005$. What linear combination of these two results would likely give a more accurate approximation to f' and why? [Hint: Richardson extrapolation.]

We can write

$$\begin{aligned} A_{.01} &= f'(x) + c(.01)^2 + d(.01)^3 \\ A_{.005} &= f'(x) + c(.005)^2 + d(.005)^3, \end{aligned}$$

for certain constants c and d. We can eliminate the $O(h^2)$ error terms by taking $\frac{4}{3}A_{.005} - \frac{1}{3}A_{.01}$. Then

$$\frac{4}{3}A_{.005} - \frac{1}{3}A_{.01} = f'(x) - \frac{4}{3}d(.005)^3.$$

7. Consider the predator-prey equations ($r(t)$ is the prey population, $s(t)$ is the predator population):

$$\begin{aligned} r' &= (2-s)r \\ s' &= (r-2)s \end{aligned}$$

Starting with $r_0 = 2$ and $s_0 = 1$, determine r_1 and s_1:

(a) Using Euler's method with stepsize $h = 0.1$. [Recall that Euler's method for solving $\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t))$ is: $\mathbf{y_{k+1}} = \mathbf{y_k} + h\mathbf{f}(t_k, \mathbf{y_k})$, $k = 0, 1, \ldots$.]

$$\begin{pmatrix} r_1 \\ s_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0.1 \begin{pmatrix} (2-1)\cdot 2 \\ (2-2)\cdot 1 \end{pmatrix} = \begin{pmatrix} 2.2 \\ 1 \end{pmatrix}.$$

(b) Using Heun's method with stepsize $h = 0.1$. [Recall that Heun's method for solving $\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t))$ is: $\mathbf{y_{k+1}} = \mathbf{y_k} + (h/2)[\mathbf{F_1} + \mathbf{F_2}]$, where $\mathbf{F_1} = \mathbf{f}(t_k, \mathbf{y_k})$ and $\mathbf{F_2} = \mathbf{f}(t_{k+1}, \mathbf{y_k} + h\mathbf{f}(t_k, \mathbf{y_k})$.]

$$F_1 = \begin{pmatrix} (2-1)\cdot 2 \\ (2-2)\cdot 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} (2-1)\cdot 2.2 \\ (2.2-2)\cdot 1 \end{pmatrix} = \begin{pmatrix} 2.2 \\ 0.2 \end{pmatrix}.$$

Therefore

$$\begin{pmatrix} r_1 \\ s_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + .05 \begin{pmatrix} 4.2 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 2.21 \\ 1.01 \end{pmatrix}.$$

8. Consider the one-step method

$$y_{k+1} = y_k + h[\theta f(t_k, y_k) + (1-\theta) f(t_{k+1}, y_{k+1})],$$

where $\theta \in [0,1]$ is given. Note that this method is *explicit* if $\theta = 1$ and otherwise it is *implicit*. Show that the local truncation error,

$$\frac{y(t_{k+1}) - y(t_k)}{h} - [\theta f(t_k, y(t_k)) + (1-\theta) f(t_{k+1}, y(t_{k+1}))],$$

where y is a true solution of the equation $y' = f(t,y)$, is $O(h^2)$ if $\theta = 1/2$ and otherwise is $O(h)$.

Expand $y(t_{k+1})$ about t_k:

$$y(t_{k+1}) = y(t_k) + hf(t_k, y(t_k)) + \frac{h^2}{2} y''(t_k) + O(h^3).$$

Subtract $y(t_k)$ from each side and divide by h to find:

$$\frac{y(t_{k+1}) - y(t_k)}{h} = f(t_k, y(t_k)) + \frac{h}{2} y''(t_k) + O(h^2). \qquad (1)$$

Now expand $y(t_k)$ about t_{k+1}:

$$y(t_k) = y(t_{k+1}) - hf(t_{k+1}, y(t_{k+1})) + \frac{h^2}{2} y''(t_{k+1}) + O(h^3).$$

Subtract $y(t_{k+1})$ from each side and divide by $-h$ to find:

$$\frac{y(t_{k+1}) - y(t_k)}{h} = f(t_{k+1}, y(t_{k+1})) - \frac{h}{2} y''(t_{k+1}) + O(h^2). \tag{2}$$

Add θ times expression (1) to $(1-\theta)$ times expression (2) to find:

$$\frac{y(t_{k+1}) - y(t_k)}{h} = \theta f(t_k, y(t_k)) + (1-\theta) f(t_{k+1}, y(t_{k+1})) + \frac{h}{2}[\theta y''(t_k) - (1-\theta) y''(t_{k+1})] + O(h^2).$$

If $\theta \neq (1-\theta)$; i.e., if $\theta \neq 1/2$, then the factor $[\theta y''(t_k) - (1-\theta) y''(t_{k+1})]$ is $O(1)$, and hence the local truncation error is $O(h)$. If $\theta = 1/2$, however, then this factor is $\frac{1}{2}[y''(t_k) - y''(t_{k+1})]$, and since $y''(t_{k+1}) = y''(t_k) + O(h)$, this factor is $O(h)$ and since it is multiplied by $h/2$ in the above expression, the local truncation error is $O(h^2)$.