This is a sample paper for 2526-MA385. It is similar to the final Semester 1 exam paper in the following ways:

- It features 5 questions; all to be attempted.
- Questions 1 and 2 are based on material from Section 1 (may have some over-lapping content. E.g., Newton's Method, or FPI could feature on both).
- Questions 3, 4 and 5 are based on Sections 2, 3, and 4 respectively with minimal overlap (and only in so far as Sections 3 and 4 overlap a little)
- Questions feature a mixture of definitions, theory and calculations.
- The questions on the exam will, of course, be different. However, if you can attempt this paper unseen, you are well prepared.

. . . .

Q1. Suppose we wish to find $\tau \in [a, b]$ such that $f(\tau) = 0$ for some nonlinear function $f(x)$.

(a) State the **Secant Method** for this problem. Provide a justification for it.

> **Answer: Derivation:** The derivation on Slide 5 of Section 1.3 will suffice. However, there are other acceptable answers, including as a relaxation method (weighted average of previous two guesses) or a discrete Newton's Method, with $f'(x_k)$ approximated as $\big(f(x_k) - f(x_{k-1})\big)/(x_k - x_{k-1})$.

(b) Suppose that $f(x) = 2x^2 - 5$. Show that $f(x) = 0$ has a solution in $[1, 2]$.

> **Answer:** Since $f(1) = -3$ and $f(2) = 3$, by the Intermediate Value theorem, there is some $x \in (-3, 3)$ such that $f(x) = 0$.

(c) Taking $x_0 = 1$ and $x_1 = 2$, carry out **three** iterations of the Secant Method to estimate the solution to $2x^3 - \mathbf{5} = 0$. Show your calculations to 4 decimal places.

> **Answer:** Computations give $x_2 = 1.5$, $x_3 = 1.5714$ and $x_4 = 1.5814$.

. . . .

Q2. (a) What does it mean for a function, g, to be a contraction on an interval $[a, b]$?

> **Answer:** It means that $a \leq g(x) \leq b$ for all $x \in [a, b]$. Furthermore, there is constant $L \in [0, 1)$ such that $|g(\alpha) - g(\beta)| \leq L|\alpha - \beta|$ for all $\alpha, \beta \in [a, b]$.

(b) Suppose that we have a fixed point iteration (FPI) method $x_{k+1} = g(x_k)$, and that g is known to be a contraction, with a fixed point τ. Show that the sequence generated by the method, $\{x_0, x_1, x_2, \dots\}$ converges *at least linearly* to τ.

> **Answer:** First note that a method converges with at least order q if there is a constant $\mu \geq 0$ such that $\lim_{k\to\infty} \frac{|\tau - x_{k+1}|}{|\tau - x_k|^q} = \mu$. (For $q = 1$ we also need that $\mu < 1$). Now we use that $g(\tau) = \tau$ and $x_{k+1} = g(x_k)$, and that it is a contraction to show that $|\tau - x_{k+1}| = |g(\tau) - g(x_k)| \leq L|\tau - x_k|$ for all k.

(c) Suppose that we want to solve $2x^2 - 5 = 0$ using FPI, in order to approximate $\tau = \sqrt{10}/2$. That is, we choose a function $g = g(x)$, and initial guess $x_0 \in [1, 2]$, and set $x_{k+1} = g(x_k)$ for $k = 0, 1, 2, \dots$. Consider the following functions:

$$g_1(x) = 2x^2 + x - 5, \qquad g_2(x) = x/2 + 5/(4x), \qquad g_3(x) = x^2/5 - 1/2.$$

For each of these, determine whether or not it is a suitable choice of g in the FPI.

> **Answer:**
>
> (i) g_1 is not a suitable choice: it is not a contraction on $[1, 3]$ since (for example) $g(1) = -2$ thus we don't have that $1 \leq g(x) \leq 2$ for $x \in [1, 2]$.
>
> (ii) This is a suitable choice.
>
> - First we note that $g(\sqrt{10}/2) = \sqrt{10}/2$.
> - Next note that $g(1) = 7/4$, $g(2) = 13/8$. Also $g'(x) = 1/2 - (5/4)x^2$, so it has a critical point at $x = \sqrt{10}/2$, at which $g_2'(\sqrt{10}/2) = 0$. So it is clear that $1 \leq g_2(x) \leq 2$ for all $x \in [1, 2]$.
> - Finally, can observe that $|g'(x)| < 1$ for $x \in [1, 2]$.
>
> (iii) g_3 is not suitable: it does not have $\sqrt{10}/2$ as a fixed point.

(d) Show that Newton's method for solving $f(x) = 0$ can be considered as a FPI method. What FPI method does it yield when we use it to solve $f(x) = 2x^2 - 5 = 0$?

> **Answer:** If we take $g(x) = x - f(x)/f'(x)$, then $g(\tau) = \tau - f(\tau)/f'(\tau) = \tau$ since $f(\tau) = 0$.
> For $f(x) = 2x^2 - 5$, we get $g(x) = x/2 + 5/(4x)$ (which, completely coincidentally, is the g_2 from Part (b).

. . . .

Q3. Consider the general two-stage Runge-Kutta (RK2) method: $y_{i+1} = y_i + f\Phi(t_i, y_I; h)$, where

$$k_1 = f(t_i, y_i), \qquad\qquad k_2 = f(t_i + \alpha h, y_i + \beta h k_1)$$

and

$$\Phi(t_i, y_i; h) = ak_1 + bk_2.$$

For a specific method we can take $\mathbf{a = 1/4}$.

(a) What does it mean for a one-step method to be *consistent*? Determine the value of $\mathbf{b}$ for the method to be consistent.

> **Answer:** A one-step method is consistent if $\Phi(t_i, y_i; 0) = f(t_i, y_i)$. (That is, if we let $h \to 0$, we the method converges to the ODE). Setting $h = 0$ above we get $k_1 = k_2 = f(t_i, y_i)$. Since $\Phi(t_i, y_i; h) = ak_1 + bk_2$, we need $a + b = 1$. This $b = 3/4$.

(b) Suppose y solves the initial value problem

$$y(1) = 1, \qquad y'(t) = 2t \quad \text{for} \quad t > 1.$$

Explain why the RK2 method should compute the exact solution. Use this fact to determine the value for α.

> **Answer:** The solution to this problem is $y(t) = t^2$. For RK2 methods, we can expect that the error is proportional to $y'''(t)$. Here $y'''(t) \equiv 0$, so the error is zero. Set $h = 1$. So we have $t_0 = 1$, $t_1 = 2$, $y_0 = 1$ and, since there is no error, t $y_1 = y(2) = 4$. Thus
>
> $$4 = y_0 + ak_1 + bk_2 = 1 + af(1,1) + bf(1+\alpha, 1+\beta f(1,1)) = 1 + (1/4)2 + (3/4)(2+2\alpha).$$
>
> Solve to get $\alpha = 2/3$.

(c) Suppose that we attempt to solve

$$y'(t) = \lambda y(t) \qquad y(0) = 1,$$

with a RK2 method. Use the fact that the RK2 solution should agree with the Taylor series for $y(t_{i+1})$ about t_i, up to terms of order h^2, to find a value of β.

> **Answer:** Since $y = e^{\lambda t}$, we have $y^{(n)}(t) = \lambda^n y(t)$. So a Taylor Series gives
>
> $$y(t_1) = y(t_0) + hy'(t_0) + \frac{h^2}{2}y''(t_0) + \frac{h^3}{6}y'''(\eta) \quad \eta \in [t_0, t_1].$$
>
> That is
>
> $$y(t_1) = y(t_0) + h\lambda y(t_0) + \frac{h^2\lambda^2}{2}y(t_0) + \mathcal{O}(h^2).$$
>
> For simplicity, set $h = 1$, to get $y(t_1) = y(t_0)(1 + \lambda + \lambda^2/2) + \mathcal{O}(h^3)$.
> The RK2 method (again with $h = 1$, works out as
> $y_1 = y_0 + (1/4)f(t_0, y_0) + (3/4)f(t_0 + \alpha, y_1 + \beta f(t_0, y_0))$. Then
> $y_1 = y_0 + (1/4)\lambda y_0 + (3/4)\lambda(y_0 + \beta\lambda y_0)$. That yields $y_1 = y_0(1 + \lambda + (3/4)\beta\lambda^2)$. For this to agree with the Taylor Series we need $\beta = 2/3$.

. . . .

Q4. (a) Let $L \in \mathbb{R}^{n\times n}$ be a non-singular lower triangular matrix, and $\boldsymbol{b} \in \mathbb{R}^n$ be such that $b_i = 0$ for $i = 1, \dots, k \le n$. If $\boldsymbol{y}$ solves $L\boldsymbol{y} = \boldsymbol{b}$, show that $y_i = 0$ for $i = 1, \dots, k \le n$.

Hence or otherwise, show that the inverse of a nonsingular lower triangular matrix is also lower triangular.

> **Answer:** Partition L by the first k rows and columns so $L\boldsymbol{y} = \boldsymbol{b}$ is
>
> $$\left(\begin{array}{c|c} L_1 & 0 \\ \hline C & L_2 \end{array}\right)\begin{pmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{\beta}\end{pmatrix}$$
>
> L is non-singular, so none of its diagonal entries are zero. Consequently L_1 is non-singular. Then $L_1\boldsymbol{y}_1 = \boldsymbol{0}$ gives that $\boldsymbol{y}_1 = \boldsymbol{0}$. That is, the first k rows of $\boldsymbol{y}$ are zero.
> Let $\boldsymbol{y}^{(j)}$ be column j of L^{-1}. It is the solution to $L\boldsymbol{y}^{(j)} = \boldsymbol{e}^{(j)}$ where the vector $\boldsymbol{e}^{(j)}$ is column j of the identity matrix. Because $e_i^{(j)} = 0$ for $i < k$, $y_i^{(j)} = 0$. So L^{-1} is lower triangular.

(b) Define the LU factorization of a matrix. What assumptions must be made on the matrix to ensure that such a factorization exists?

> **Answer:** The unit lower triangular matrix L and upper triangular matrix U are the LU-factorisation of A if $A = LU$. Such a factorization is possible is every leading principle $k \times k$ submatrix of A is nonsingular for $k = 1, 2, \dots, n-1$.

(c) Find the LU-factorisation of

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & -3 & 4 \end{pmatrix}.$$

Use this factorization to solve $Ax = b$, where $b = (4, -4, 0, 8)^T$.

> **Answer:** First we factorise A (which can be done by inspection, or by applying a formula).
>
> $$\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & -3 & 4 \end{pmatrix} . = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}}_{L}\underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}}_{U}$$
>
> Now solve $Ly = b$ (where b is the RHS), giving $y = (4, 0, 0, -8)^T$. Then solve $Ux = y$, to get $x = (4, 0, 0, -2)^T$.

. . . .

Q5. (a) Recall the definition of the Euclidean norm on $\mathbb{R}^n$: $\|\boldsymbol{u}\|_2 = \sqrt{\boldsymbol{u}^T\boldsymbol{u}}$. Prove the Cauchy-Schwarz inequality:

$$|\sum_{i=1}^{n} u_i v_i| \leq \|\boldsymbol{u}\|_2 \|\boldsymbol{v}\|_2 \qquad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n.$$

Hence show that $\|\cdot\|_2$ satisfies the triangle inequality.

Answer: To prove the Cauchy-Schwarz: for any $\lambda \in \mathbb{R}$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,

$$0 \leq \|\lambda \boldsymbol{u} + \boldsymbol{v}\|_2^2 = \sum_{i=1}^{n} (\lambda u_i + v_i)^2 = \lambda^2 \|u\|_2^2 + 2\lambda \sum_{i=1}^{n} u_i v_i + \|\boldsymbol{v}\|_2^2.$$

This polynomial in λ has at most one real root, so $\big(2\sum_{i=1}^n u_i v_i\big)^2 - 4\|\boldsymbol{u}\|_2^2\|\boldsymbol{v}\|_2^2 \leq 0$. Thus

$$|\sum_{i=1}^{n} u_i v_i| \leq \|\boldsymbol{u}\|_2 \|\boldsymbol{v}\|_2.$$

Clearly $\|\cdot\|_2$ satisfies (a)-(c) above. From C-S

$$0 \leq \|\boldsymbol{u}+\boldsymbol{v}\|_2^2 = \|\boldsymbol{u}\|_2^2 + 2\sum_{i=1}^{n} u_i v_i + \|\boldsymbol{v}\|_2^2 \leq \|\boldsymbol{u}\|_2^2 + 2\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2 + \|\boldsymbol{v}\|_2^2 = (\|\boldsymbol{u}\|_2 + \|\boldsymbol{v}\|_2)^2,$$

so it also satisfies the triangle inequality.

(b) Let A be any matrix in $\mathbb{R}^{n\times n}$. What are the *singular values* of A? Show that they are real and non-negative.
Define the *subordinate matrix norm* on $\mathbb{R}^{n\times n}$ associated with $\|\cdot\|_2$ and show that $\|A\|_2$ is the largest singular value of A.

Answer: The singular values of AA are the eigenvalues of $B = A^T A$. The proof that they are read and non-negative is in Section 4.2 of the notes (see Theorem 4.2.3).
Defn: Subordinate matrix norm:

$$\|A\|_2 = \max_{\boldsymbol{v}\in\mathbb{R}^n_\star} \frac{\|A\boldsymbol{v}\|_2}{\|\boldsymbol{v}\|_2}, \qquad A \in \mathbb{R}^{n\times n}, \mathbb{R}^n_\star = \mathbb{R}^n/\{\boldsymbol{0}\}.$$

See proof of Thm 4.2.4 for the rest.

(c) State the Gerschgorin First Circle Theorem, and use it to find an upper bound on $\|A\|_2$ when

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & -3 & 4 \end{pmatrix}.$$

Answer: Given a matrix $A \in \mathbb{R}^{n \times n}$, let D_i be the discs in the complex plane centre a_{ii} and radius r_i:

$$r_i = \sum_{j=1, j \neq i}^{n} |a_{ij}|.$$

Gerschgorin's First Theorem: All the eigenvalues of A are contained in the union of the Gerschgorin discs.

$$B = A^T A = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 18 & 3 & 0 \\ 1 & 3 & 11 & -12 \\ 0 & 0 & -12 & 16 \end{pmatrix}$$

So the eigenvalues are contained in the intervals $[1, 3] \cup [15, 21] \cup [-5, 27] \cup [4, 28]$. So $\|A\|_2^2$ is at most $\sqrt{28} = 5.2915$.