EAS501 Midterm Review

February 28, 2024

Contents

1 Introduction

These notes surely have some errors. If there are any discrepancies between these notes and the lecture notes, always assume the lecture notes are correct

2 Matlab

refer to UB learns.

3 Vectors

- A vector is an organized collection of number called components.
  - For physicists, a vector is an arrow with a magnitude and length.
  - For data scientists, a vector might be a collection of data.
  - For a mathematician, a vector might be an element of a vector space.
- We can column vectors:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

- and row vectors:

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

- vectors have the following operations:
  - addition
  - subtraction
  - scalar multiplication
  - cross product
  - transpose

3.1 Dot product

Definition:

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w} = |\mathbf{v}| \cdot |\mathbf{w}| \cdot \cos(\theta) = \sum_{i=1}^{n} v_i \cdot w_i$$

Example: Let $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$.

The dot product of $\mathbf{v}$ and $\mathbf{w}$ is calculated as:

$$\mathbf{v} \cdot \mathbf{w} = (1 \cdot 4) + (2 \cdot 5) + (3 \cdot 6) = 4 + 10 + 18 = 32$$

3.2 Properties

The follow properties are also true:

- Commutative Property:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$$

- Distributive Property:

$$(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{w} + \mathbf{x}) = \mathbf{u} \cdot \mathbf{w} + \mathbf{u} \cdot \mathbf{x} + \mathbf{v} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{x}$$

- Associative Property:

$$a(\mathbf{u} \cdot \mathbf{v}) = (a\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (a\mathbf{v})$$

3.3 Norms

Norms give us information about the size of a vector. Each norm gives us a different perspective on the "size" or "length".

p-norm:

$$\|\mathbf{u}\|_p = \left(\sum_{i=1}^{n} |u_i|^p \right)^{1/p}$$

2-norm (Euclidean norm):

$$\|\mathbf{u}\|_2 = \sqrt{\sum_{i=1}^{n} |u_i|^2}$$

1-norm:

$$\|\mathbf{u}\|_1 = \sum_{i=1}^{n} |u_i|$$

∞-norm:

$$\|\mathbf{u}\|_\infty = \max_i |u_i|$$

3.3.1 All norms obey the following rules

- Positive Definiteness:

$$\|\mathbf{u}\|_p \geq 0 \quad \text{and} \quad \|\mathbf{u}\|_p = 0 \text{ if and only if } \mathbf{u} = \mathbf{0}$$

- Scalar Multiplication:

$$\|k\mathbf{u}\|_p = |k| \|\mathbf{u}\|_p$$

- Triangle Inequality:

$$\|\mathbf{u} + \mathbf{v}\|_p \leq \|\mathbf{u}\|_p + \|\mathbf{v}\|_p$$

3.4 Cauch-Schwartz Inequality

Tells us the upper bound on the inner product (dot product) of two vectors.

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|$$

3.5 Linear combinations of vectors

- A linear combination of vectors is a weighted sum of the vectors.
- An example of a linear combination of vectors is given by:

$$a\mathbf{u} + b\mathbf{v} + c\mathbf{w}$$

where a,b, and c are scalars.

- If this doesn't make sense to you, pause here before you continue. This will come back a lot after the midterm.

3.6 Problems

- What does it mean if the dot product of two vectors equals 0?

4 Matrices

A matrix is a collection of 2D numbers. They can represent a collection of data, a system of linear equations, a linear mapping (more on this after the midterm), an adjacency matrix of a graph, a markov transition matrix, an epic movie from 1999, etc. etc.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \end{bmatrix}$$

- the size of a matrix is m $\times$ n where m is the number of rows and n is the number of columns. In the above example, m $=$ 4, and n $=$ 5. We write

$$A \in \mathbb{R}^{4\times 5}$$

- The transpose of matrix A is denoted as $A^\top$, which is given by:

$$A^\top = \begin{bmatrix} 1 & 6 & 11 & 16 \\ 2 & 7 & 12 & 17 \\ 3 & 8 & 13 & 18 \\ 4 & 9 & 14 & 19 \\ 5 & 10 & 15 & 20 \end{bmatrix}$$

4.1 Matrix vector product as a linear combination

Let A be an $m \times n$ matrix represented as a collection of column vectors:

$$A = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix}$$

where $\mathbf{c}_i$ represents the i-th column vector of A.

Now, let $\mathbf{x}$ be a vector of size n:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

The product of A and $\mathbf{x}$ can be expressed as:

$$A\mathbf{x} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

This multiplication yields:

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

Thus, multiplying the matrix A by the vector $\mathbf{x}$ results in a linear combination of the column vectors of A.

4.2 Operations

- Addition/ Subtraction of equally sized matrices
- Matrix Vector products
- Matrix Matrix products
- Matrix Powers

- Block Matrices
- Trace = sum of the diagnoal elements
- outer product
- matrix determinant (know up to 3x3 by hand)

4.3 Properties

1. Commutative Property: $A + B = B + A$
2. Associative Property: $(A + B) + C = A + (B + C)$
3. Identity Element: $A + \mathbf{0} = \mathbf{0} + A = A$, where $\mathbf{0}$ is the zero matrix of appropriate size.
4. Associative Property: $(AB)C = A(BC)$
5. Distributive Properties:
   - Matrix-Matrix Multiplication: $A(B + C) = AB + AC$
   - Matrix-Scalar Multiplication: $k(AB) = (kA)B = A(kB)$
6. Matrix Multiplication is generally **not** commutative: $AB \neq BA$ for most matrices A and B.
7. Matrix-Vector Multiplication: $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$
8. Scalar-Vector Multiplication: $k(A\mathbf{x}) = (kA)\mathbf{x}$

Properties of Determinants:

1. Multiplicative Property: $\det(AB) = \det(A)\det(B)$ for matrices A and B of the same size.
2. Scalar Multiplication: $\det(aA) = a^n \det(A)$ for a scalar a and an $n \times n$ matrix A.
3. Determinant of the Transpose: $\det(A^\top) = \det(A)$ for any square matrix A.
4. The determinant of a mtrix is equal to the area parallelogram defined by the columns of the matrix.

4.4 Matrix Inverse

$A^{-1} = \frac{1}{\det(A)}\mathrm{adj}(A)$

4.5 Problems

- outer product of two vectors
- trace of a matrix
- matrix determinant

5 Matrix Norms, Graphs, and Markov Chains

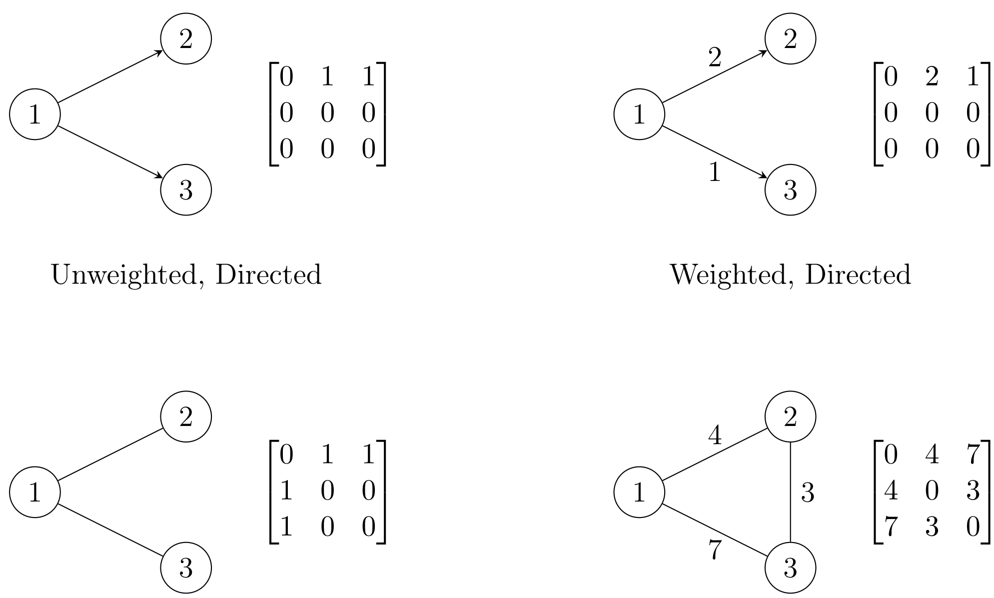
5.1 Matrix Norms

Matrix norms give us information about the "size" or "extent" of a matrix.

$$
\begin{aligned}
\text{1-norm (largest column 1-norm)} : \quad & \|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \\
\text{Infinite norm (largest row 1-norm)} : \quad & \|A\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}| \\
\text{Frobenius norm (square root of sum of elements squared)} : \quad & \|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \\
\text{2-norm (spectral norm)} : \quad & \|A\|_2 = \sigma_{\max}(A) \leq \|A\|_F
\end{aligned}
$$

5.2 Graphs

- Graph - a collection of vertices connected by edges.
- Adjacency matrix - matrix representation of a graph including connections, weights, and direction.
- Path - finite sequence of edges that connects two vertices
- Length - number of edges in a path.
- A^n (i,j) = number of paths of length n between i and j



$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Unweighted, Directed

$$\begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Weighted, Directed

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

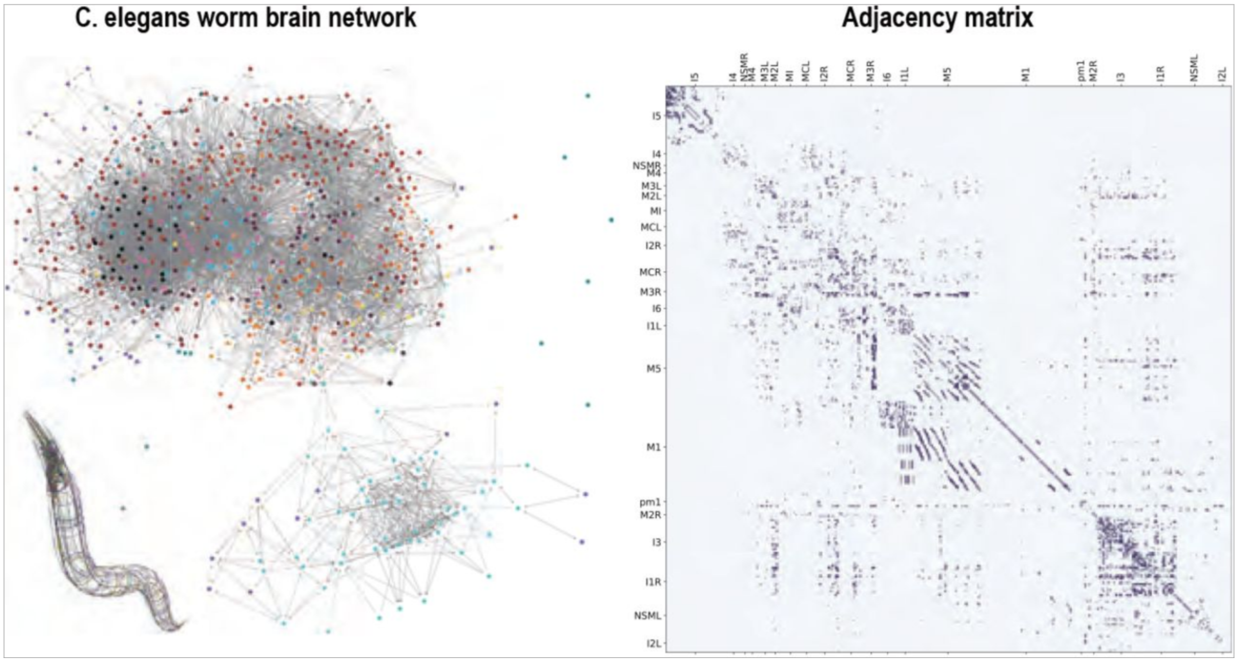
Unweighted, Undirected

$$\begin{bmatrix} 0 & 4 & 7 \\ 4 & 0 & 3 \\ 7 & 3 & 0 \end{bmatrix}$$

Weighted, Undirected

Figure 1: Graphs and their Adjacency Matrices

- graphs can get complex



- mathematicians have developed powerful algorithms for extracting information from these networks

5.3 Markov Chains

- the probability vector tells us about the current state of the system.

$$p_n = \begin{bmatrix} 0.8 \\ 0.1 \\ 0.1 \end{bmatrix}$$

- The markov transition matrix tells us how this system will change in one iteration.

$$M = \begin{bmatrix} 0.8 & 0.1 & 0.4 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}$$

- Multiplying the transition matrix by the probability vector gives us the

$$Mp_n = \begin{bmatrix} 0.8 & 0.1 & 0.4 \\ 0.1 & 0.6 & 0.2 \\ 0.1 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} 0.8 \\ 0.1 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.8(0.8) + 0.1(0.1) + 0.4(0.1) \\ 0.1(0.8) + 0.6(0.1) + 0.2(0.1) \\ 0.1(0.8) + 0.3(0.1) + 0.4(0.1) \end{bmatrix} = \begin{bmatrix} 0.73 \\ 0.15 \\ 0.12 \end{bmatrix} = p_{n+1}$$

- So in general we have $p_n = M^n p_0$ (derive this yourself)
- Markov chain transition matrices are **stochastic** matrix:
  - square
  - all entries are non-negative
  - all columns add to 1
- important properties include:
  - $\lim_{n\to\infty} M^n = M_\infty$
  - All values of M_∞ are strictly positive
  - All columns of M_∞ are the same
  - p_∞ is a column of M_∞ (it doesn't matter the original p_0)

- if the Markov chain transition matrix is also **regular** (it has strictly positive values for some M^{k}) then, it will reach an equilibrium p_{∞}, otherwise, it will not.

5.4 Problems

6 Linear Systems and RREF

- a system of linear equations is when we have 2 or more linear equations that are describing the same system.
- Solving a system of linear equations can lead to one of three outcomes:
  - one unique solution
  - infinite number of solutions
  - no solutions
- How can we find a solution to a system of linear equations? One way is RREF.

6.1 RREF

System of linear equations:

$$\begin{aligned} 2x + 3y &= 5 \\ 4x - y &= 3 \end{aligned}$$

Matrix-vector multiplication form:

$$\begin{bmatrix} 2 & 3 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Augmented matrix form:

$$\left[\begin{array}{cc|c} 2 & 3 & 5 \\ 4 & -1 & 3 \end{array}\right]$$

With RREF we can do the following to the augmented matrix form:

- swap rows
- multiply a row by a scalar
- add rows

Convince yourself that these are all legal algebraic operations on the system of linear equations.

When a matrix is in RREF form:

- It is in echelon form (echelon = "steps")
- Each pivot is 1
- Each column with a pivot only has a non-zero number at that pivot

RREF tells us if a system of linear equations has:

- one unique solution - number of pivots equals number of rows AND the matrix is square
- infinite solutions - number of pivots is less than the number of rows OR columns, AND all rows are consistant
- No solution - inconsistent row

If one of these statements is true then ALL are true (and vice versa)

- A is invertable (A^{-1} exists)
- $\det(A) \neq 0$
- rref(A) has n-pivots
- rref(A) = I
- Ax = b has a unique solution for all $b \in R^{n \times 1}$

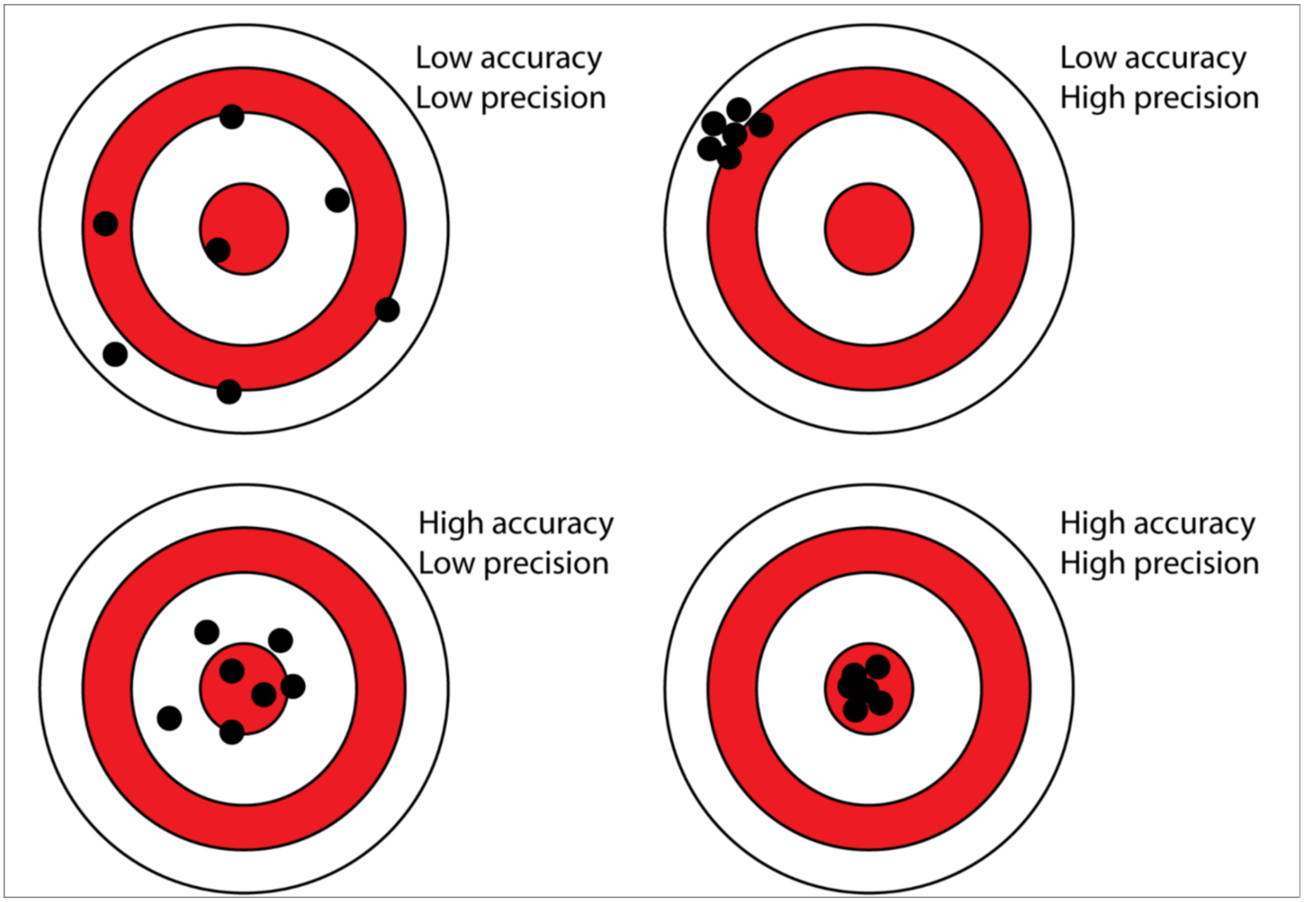
Row Operation 1:	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 4 \end{matrix}$	add -**1** times the 2nd row to the 3rd row	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{matrix}$
Row Operation 2:	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{matrix}$	multiply the 3rd row by **1/3**	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{matrix}$
Row Operation 3:	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{matrix}$	add -**1** times the 3rd row to the 2nd row	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$
Row Operation 4:	$\begin{matrix} 1 & 3 & 8 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$	add -**3** times the 2nd row to the 1st row	$\begin{matrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}$

6.2 Problems

- why do we pivot? (Gaussian elimination with pivoting?)

7 Interpolation

7.1 Accuracy vs Precision



- On your own, study the following:
  - significant figure
  - single precision (32-bit, 7 sig figs)
  - double precision (64-bit, 15 sig figs)
  - rounding
  - chopping
  - round-off errors
  - condition number

7.2 Interpolation

Here's the idea. We have a data set (temperature values in buffalo at different times, for example) but the data is not complete. We want to come up with a good estimate for the missing values based on the known values. There are two main types of interpolation, **global interpolation** and **piecewise interpolation**:

7.2.1 Global interpolation

- Use all of the data to construct one function over the entire domain.

Example: **Polynomial interpolation**

- Construct a polynomial function that passes through a given set of data points.

- We have a set of n data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$,
- We want to find a polynomial function $f(x)$ such that $f(x_i) = y_i$ for $i = 1, 2, \ldots, n$.
- The general form of a polynomial of degree $n - 1$ is given by:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$$

- To find the coefficients $a_0, a_1, \ldots, a_{n-1}$, we substitute each data point (x_i, y_i) into the polynomial:

$$\begin{aligned} f(x_1) &= a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_{n-1} x_1^{n-1} = y_1 \\ f(x_2) &= a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_{n-1} x_2^{n-1} = y_2 \\ &\vdots \\ f(x_n) &= a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_{n-1} x_n^{n-1} = y_n \end{aligned}$$

- These equations form a system of linear equations
- We can express it in matrix form as $V\mathbf{a} = \mathbf{y}$, where V is the Vandermonde matrix, $\mathbf{a}$ is the column of unknown vector of coefficients $[a_0, a_1, \ldots, a_{n-1}]^T$, and $\mathbf{y}$ is the column vector of y-values $[y_1, y_2, \ldots, y_n]^T$.
- Vandermonde matrix V is constructed by arranging the powers of x_i as columns:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- The solution to the system $V\mathbf{a} = \mathbf{y}$ can be found by solving for $\mathbf{a}$, which gives us the coefficients of the polynomial function $f(x)$.
- However, V has a huge condition number. Its not stable! So maybe this isn't the best way of doing it...

Example: **Lagrange Interpolation**:

- Instead of simple polynomials, we can construct polynomials with certain constraints that will be better suited for interpolation.

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}$$

- Properties of Lagrange Polynomials:

1.
$$\sum_{i=1}^{n} L_i(x) = 1 \quad \text{for all } x$$

2.
$$L_i(x_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

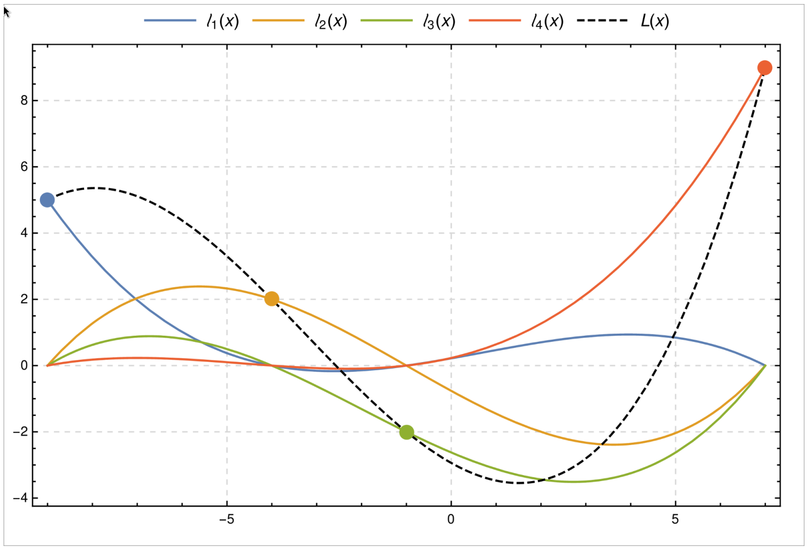
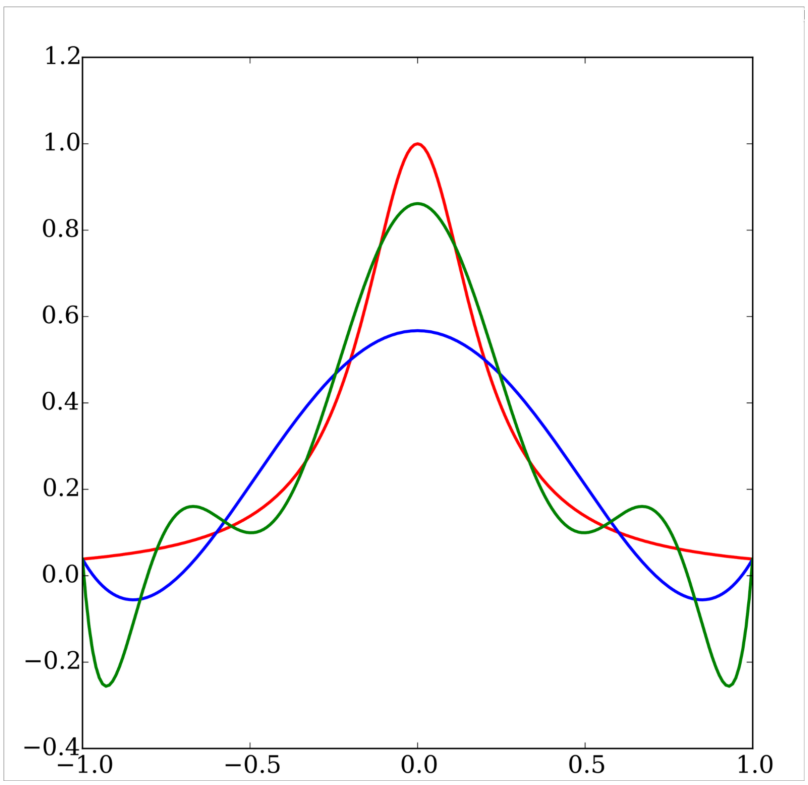


Figure 2: The first 4 lagrange polynomials, interpolating a n=4 data set

7.2.2 Runge's Phenomena and Chebyshev points

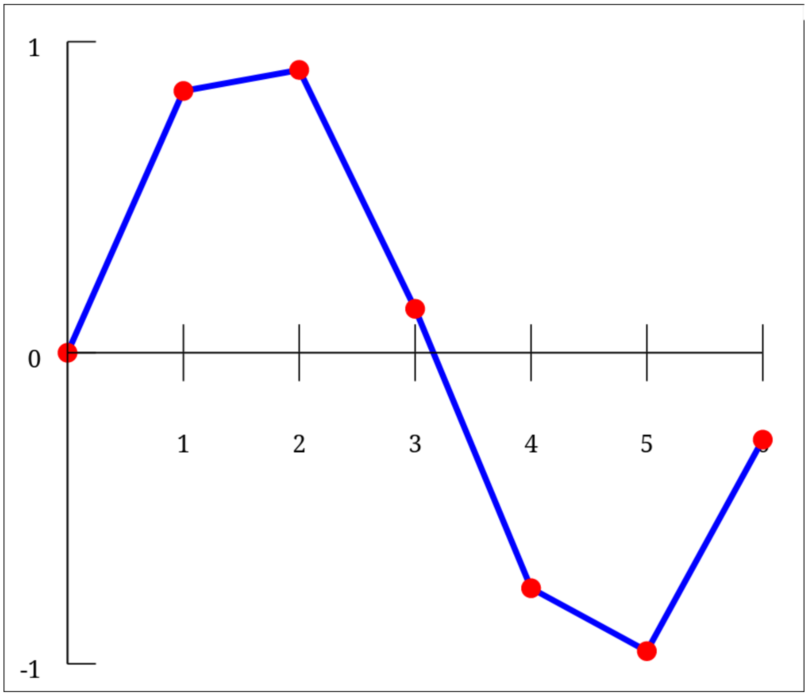


7.2.3 Piecewise interpolation

- Use one equation between each set of data points
- Each interpolating function between two data points is called a **spline**.
- Given n-data points (also called knots), we need n-1 splines
- In this class, we'll study:
  - linear splines
  - cubic splines

- hermite interpolation
- Radial basis interpolation

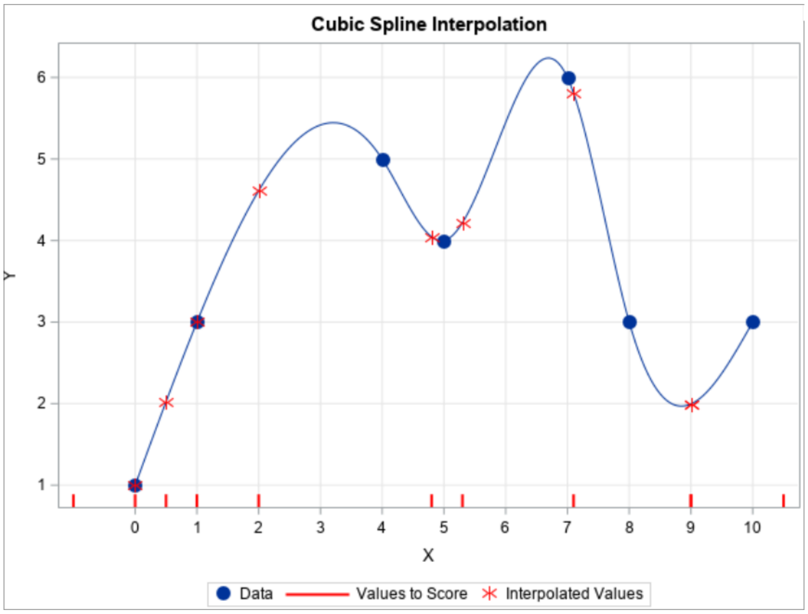
- The idea is always the same:
  1. Assume a certain form of the spline function
  2. Apply conditions that each spline must adhere to. For example:
     - splines must return the same value at the same point
     - splines must be continuous
  3. Use these conditions to create a solvable system of linear equations.
  4. Solve for the unknowns to find the equations for each spline.
- Example: **linear splines**:



1. Assume splines of the form $f_i(x) = a_i + b_i(x - x_i)$
2. Apply the following condition:
   - Splines must return the correct y value at each data point
   - Linear splines must be C^0 continuous at each data point
3. Use this condition to create a set of linear equations
4. Solve

(Example in class)

- Example 2: **cubic splines**



1. Assume splines of the form $f_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3$
   - 4 unknowns for each spline, we are going to need to impose some conditions to solve this
2. Apply the following conditions
   - Splines must return the correct y value at each knot
   - Splines must be C^0 continuous at common knots
   - Splines must be C^1 continuous at common knots
   - Splines must be C^2 continuous at common knots
   - **we still need two more equations to solve this system!** Pick one of the following:
     - natural cubic splines
     - Clamped Cublic splines
     - "not-a-knot"
3. Use this condition to create a set of linear equations
4. Solve

- Example 2: **Hermite interpolation** (we are given data points, and their derivative at a point)
  1. Assume splines of the form $f_i(x) = p_i\left(\frac{x-x_i}{h_i}\right)$ where

$$p_i(t) = (2t^3 - 2t^2 + 1)y_i + (t^3 - 2t^2 + t)h_i y_i^{'} + (-2t^3 + 3t^2)y_{i+1} + (t^3 - t^2)h_i y_{i+1}^{'}$$

  2. Apply the following conditions
     - $p_i(1) = p_{i+1}(0)$
     - $p_i^{'}(1) = p_{i+1}^{'}(0)$
     - ...
  3. Create a system of linear equations
  4. Solve
- Example 2: **Radial Basis Interpolation** (functions depend only on radial distance between points)

1. Assume splines of the form $s(x) = \sum_{i=1}^{n} w_i \phi(\|x - x_i\|)$, where $s(x)$ is the interpolating function, x is the input vector, x_i are the data points, w_i are the weights to be determined, $\phi(\|x - x_i\|)$ is the radial basis function, and n is the number of data points.
2. Apply the following conditions
   - $\mathrm{s(x_i) = f_i}$
3. Create a system of linear equations
4. Solve

With radial basis interpolation we can choose our kernel ϕ, based on the problem:

- Gaussian
- Multiquadratic
- Inverse Multiquadratic
- Polyharmonic Spline

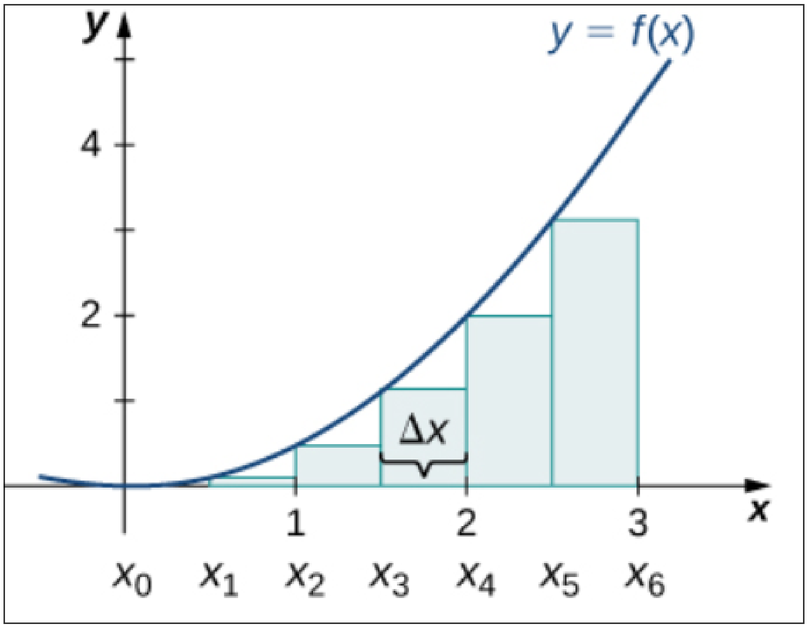
7.3 Problems

- How many sig figs in 0.001234?
- Round 0.00024382 to 2 sig figs with chopping

8 Integration

- You've learned to integrate functions in you calculus class.
- What if we want to integrate functions that don't have a nice analytic solutions?
- We use numerical integration methods.

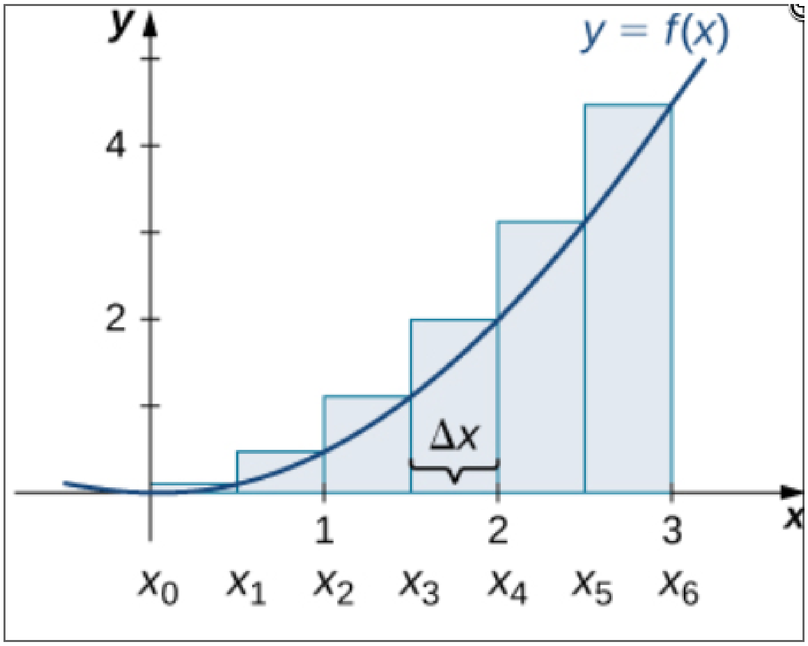
8.1 Left point approximation



$$\int_b^a f(x)dx \approx \sum_{i=1}^{n-1} f(x_i)h_i$$

- Where $h_i = x_i - x_{i-1}$
- x_i is the left data point
- local error: $O(h^2)$
- global error: $O(h)$

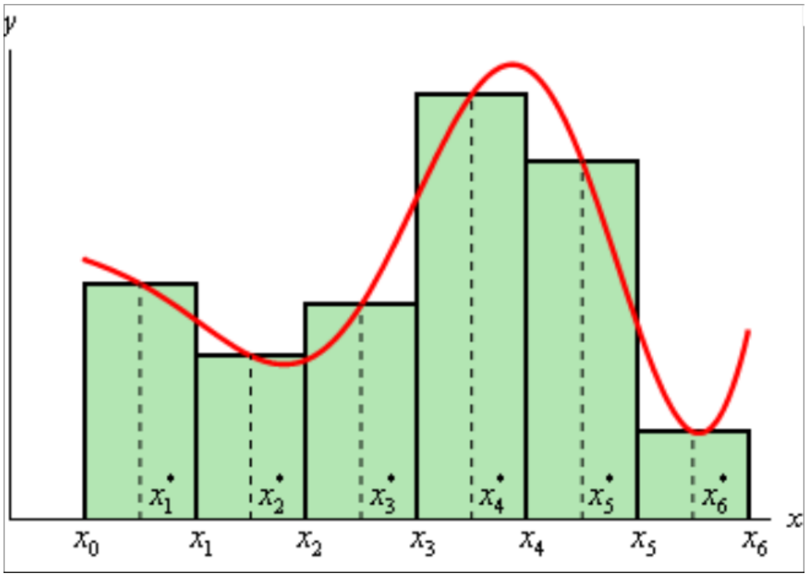
8.2 Right Point approximation



$$\int_b^a f(x)dx \approx \sum_{i=1}^{n-1} f(x_i)h_i$$

- Where $h_i = x_i - x_{i-1}$
- x_i is the right data point
- - local error: $O(h^2)$
- global error: $O(h)$

8.3 Midpoint Rule

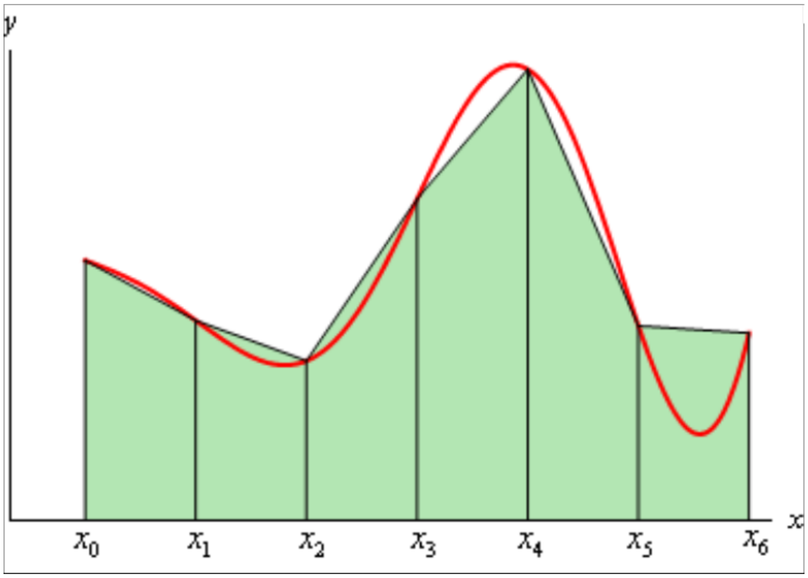


$$\int_b^a f(x)dx \approx \sum_{i=1}^{n} f(\overline{x_i})h_i$$

- Where $\overline{x_i} = \frac{x_{i-1}+x_i}{2}$

- $h_i = x_i - x_{i-1}$
- local error: $O(h^3)$
- global error: $O(h^2)$

8.4 Trapezoid Rule



$$\int_b^a f(x)dx \approx \sum_{i=1}^{n} \frac{1}{2}(f(x_{i-1}) + f(x_i))h_i$$

- $h_i = x_i - x_{i-1}$
- local error: $O(h^3)$
- global error: $O(h^2)$

8.5 Simpson's Rule

$$\int_b^a f(x)dx \approx \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{6}\left[f(x_{i-1}) + 4f\left(\frac{x_i + x_{i-1}}{2}\right) + f(x_i)\right]$$

- local error: $O(h^5)$
- global error: $O(h^4)$

8.6 Gaussian Quadrature

$$\int_{-1}^{1} f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i)$$

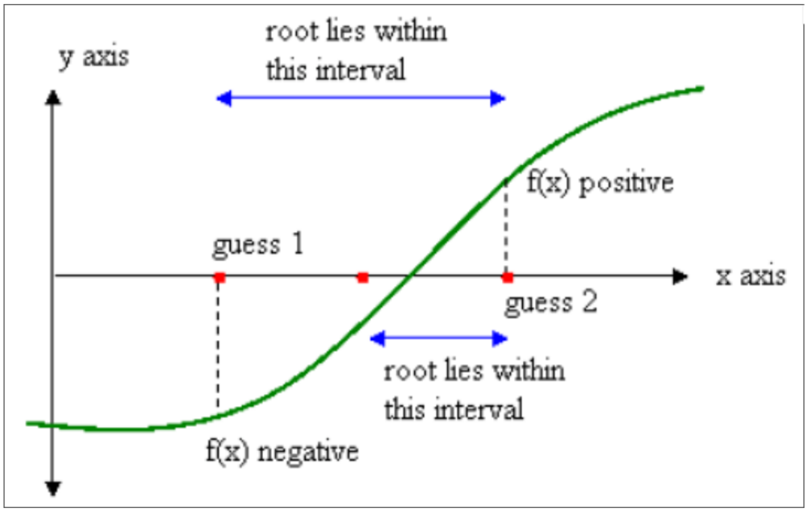
- Quadrature rule that gives an exact result for $\int_{-1}^{1} f(x)dx$ where $f(x)$ is a polynomial of degree 2n-1 or less.
- the weights, w_i are predefined (look them up wikipedia)
- So, a 7 point Gaussian quadrature can integrate a polynomial of degree $2(7) - 1 = 13$ or less, **exactly** over the interval [-1, 1]

8.7 Problems

9 Root Finding

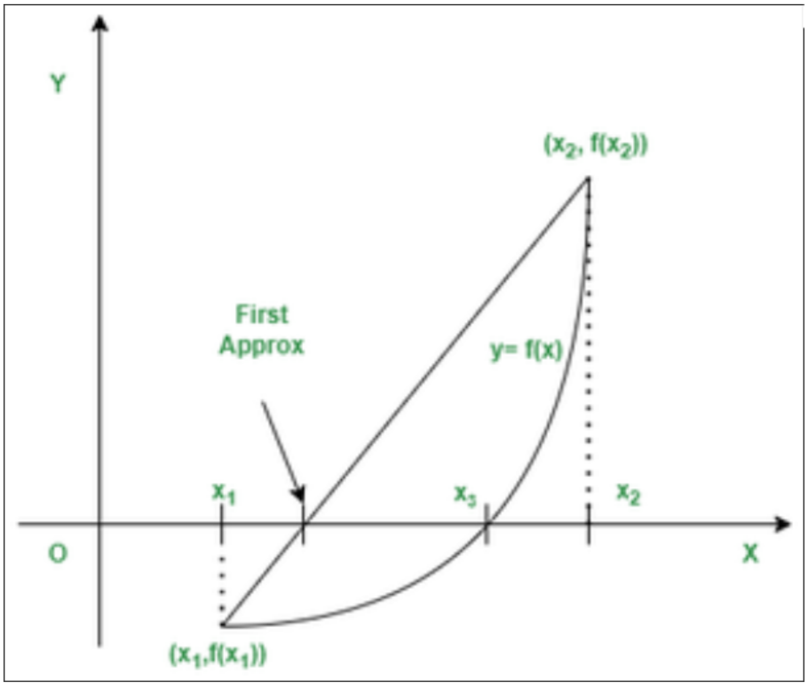
- We are looking for places where a function, f, has a root ($f(x) = 0$)
- solving an equation $f(x) = g(x)$ is the same as finding the roots of the function $h(x) = f(x) - g(x)$
- let ϵ be a small number.
  - convergence criteria: our guess for the root, $f(x_i)$ is smaller than ϵ
- let δ be a comparitivly large number
  - divergence criteria: the difference between two successive guesses, $f(x_1)$ and $f(x_2)$ is larger than δ

9.1 Bisection Method



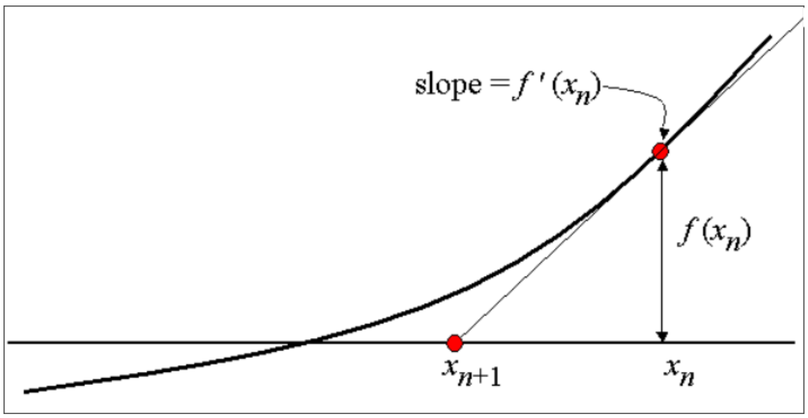
- solve for $f(x) = 0$ in $x \in [a, b]$
- Guess two values for x (a and b) such that $f(a)f(b) < 0$ (if there is a sign change, meaning f(x) must cross 0)
- let the midpoint be $m = \frac{a+b}{2}$
- Algorithm:
  - check $|b - a| < \delta$, the divergence criteria, return m
  - check $|f(m)| < \epsilon$, the convergence criteria, return m
  - if $f(a)f(m) < 0$, set new interval to $[a, m]$ and repeat
  - if $f(b)f(m) < 0$, set new interval to $[m, b]$

9.2 Regula Falsi Method



- Again, guess two values for x (a and b) such that $f(a)f(b) < 0$ (if there is a sign change, meaning f(x) must cross 0)
- Draw a line connecting $(a, f(a)), (b, f(b))$: $y(s) = f(a) + \frac{f(b)-f(a)}{b-a}(s-a)$
- Solve for where this line crosses the axis (ie. $y(s) = 0$): $s = \frac{bf(a)-af(b)}{f(a)-f(b)}$
- check $|b-a| < \delta$, the divergence criteria, return s
- check $|f(m)| < \epsilon$, the convergence criteria, return s
- if $f(a)f(s) < 0$, set new interval to $[a, s]$ and repeat
- if $f(b)f(s) < 0$, set new interval to $[s, b]$

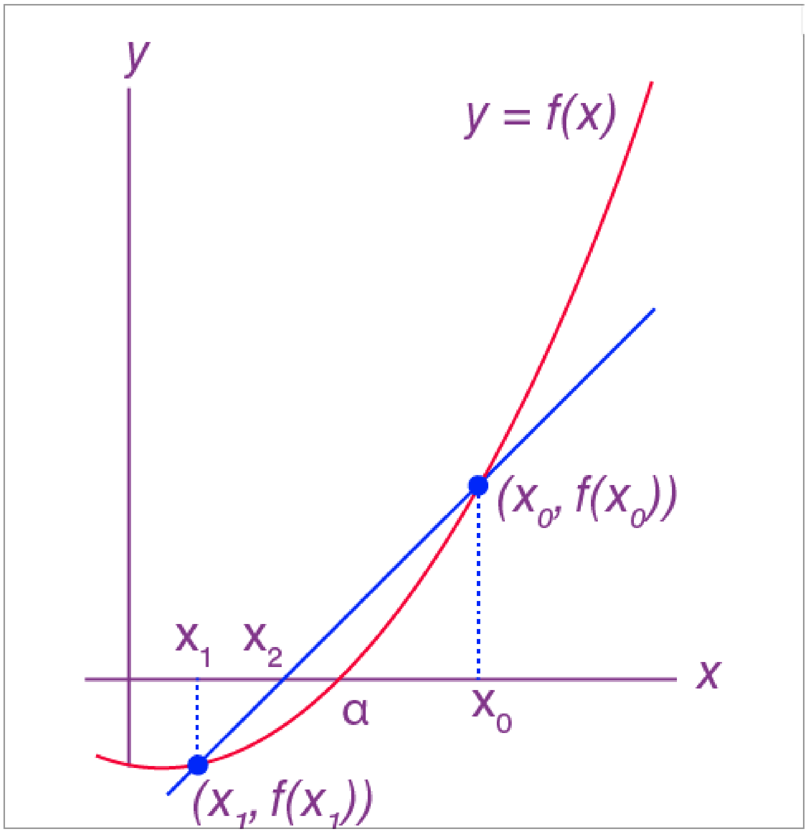
9.3 Newton Rhapson Method



$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

- might diverge. When?
- is divergence dependent on the initial guess?

9.4 Secant Method



$$x_{n+1} = x_n - f(x_n)\frac{f(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$$

9.5 Fixed Point Method

- a fixed point of a function $f(x)$ is a point x_0 such that $f(x_0) = x_0$
- split f(x) into a linear part, ax and a non-linear part $g(x)$
- Example from homework

10 Nonlinear Systems

Lets do root finding for n-nonlinear equations with n unknowns

$$f_1(x_1, x_2, \ldots, x_n) = 0 \tag{1}$$
$$f_2(x_1, x_2, \ldots, x_n) = 0 \tag{2}$$
$$\vdots \tag{3}$$
$$f_n(x_1, x_2, \ldots, x_n) = 0 \tag{4}$$

We can write it as a vector equation:

$$\boldsymbol{f}(\mathbf{x}) = \mathbf{0}$$

10.1 Fixed-point

- like before, lets split it into a nonlinear and a linear part $

$$\boldsymbol{f}(\mathbf{x}) = \boldsymbol{g}(\mathbf{x}) + \boldsymbol{A}\mathbf{x} = \mathbf{0}$$

- Now lets do the rest by hand with an example...

10.2 Multidimensional Newton-Rhaphson

First lets understand the Jacobian:

10.2.1 Jacobian

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Here, $f_1, f_2, \ldots, f_m$ are the components of the vector-valued function $\boldsymbol{f}(\mathbf{x})$, and $x_1, x_2, \ldots, x_n$ are the variables.

10.2.2 formula

$$x_{i+1} = x_i - (J_i)^{-1} \cdot f(x_i)$$