Preparing to Mid-Semester Test

COMPSCI 369 S1C

April 27, 2011

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2f(x)}{d^2x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2 f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2 f(x)}{d^2 x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2f(x)}{d^2x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2f(x)}{d^2x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2 f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2 f(x)}{d^2 x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Taylor Series Expansion

Write down the second-degree Taylor polynomial $\widehat{f}_2(x;1)$ to approximate the function $f(x) = \frac{1}{x}$ in neighbourhood of $x = 1$ and find the approximation error $e = |\widehat{f}_2(0.5;1) - f(0.5)|$ at $x = 0.5$

Hint: $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$

- Second-degree Taylor series approximation around $x = 1$:

$$\widehat{f}_2(x;1) = f(1) + (x-1)\left.\frac{df(x)}{dx}\right|_{x=1} + \frac{(x-1)^2}{2}\left.\frac{d^2f(x)}{dx^2}\right|_{x=1}$$

- For $f(x) = \frac{1}{x}$, $f'(x) \equiv \frac{df(x)}{dx} = -\frac{1}{x^2}$ and $f''(x) \equiv \frac{d^2f(x)}{d^2x} = \frac{2}{x^3}$
- At point $x = 1$, $f(1) = 1$; $f'(1) = -1$, and $f''(1) = 2$
- $\widehat{f}_2(x;1) = 1 - (x-1) + (x-1)^2 = x^2 - 3x + 3$
- Error at $x = 0.5$:
  $\left|\frac{1}{0.5} - (0.5)^2 + 3 \cdot 0.5 - 3\right| = |2 - 0.25 + 1.5 - 3| = 0.25$

Newton's Root Finder

Explain the basic idea of Newton's method for finding roots of a function $f(x)$ and show the first step of searching for the root of $f(x) = x^3 - x = 0$, starting from $x_0 = 2$

- The method is based on computing each next approximation x_{n+1} of the goal root using the first-degree Taylor series expansion of $f(x)$ around the current point x_n:

$$f(x) \approx \widehat{f}_1(x; x_n) = f(x_n) + (x - x_n)f'(x_n)$$

- Next approximation x_{n+1} follows from $\widehat{f}_1(x_{n+1}; x_n) = f(x_n) + (x_{n+1} - x_n)f'(x_n) = 0$:

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

- For $f(x) = x^3 - x = 0$, $f'(x) = 3x^2 - 1$; $f(2) = 6$, and $f'(2) = 11$. Therefore, $x_1 = 2 - \frac{6}{11} \approx 1.45$

Newton's Root Finder

Explain the basic idea of Newton's method for finding roots of a function $f(x)$ and show the first step of searching for the root of $f(x) = x^3 - x = 0$, starting from $x_0 = 2$

- The method is based on computing each next approximation x_{n+1} of the goal root using the first-degree Taylor series expansion of $f(x)$ around the current point x_n:

$$f(x) \approx \widehat{f}_1(x; x_n) = f(x_n) + (x - x_n)f'(x_n)$$

- Next approximation x_{n+1} follows from $\widehat{f}_1(x_{n+1}; x_n) = f(x_n) + (x_{n+1} - x_n)f'(x_n) = 0$:

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

- For $f(x) = x^3 - x = 0$, $f'(x) = 3x^2 - 1$; $f(2) = 6$, and $f'(2) = 11$. Therefore, $x_1 = 2 - \frac{6}{11} \approx 1.45$

Newton's Root Finder

Explain the basic idea of Newton's method for finding roots of a function $f(x)$ and show the first step of searching for the root of $f(x) = x^3 - x = 0$, starting from $x_0 = 2$

- The method is based on computing each next approximation x_{n+1} of the goal root using the first-degree Taylor series expansion of $f(x)$ around the current point x_n:

$$f(x) \approx \widehat{f}_1(x; x_n) = f(x_n) + (x - x_n)f'(x_n)$$

- Next approximation x_{n+1} follows from $\widehat{f}_1(x_{n+1}; x_n) = f(x_n) + (x_{n+1} - x_n)f'(x_n) = 0$:

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

- For $f(x) = x^3 - x = 0$, $f'(x) = 3x^2 - 1$; $f(2) = 6$, and $f'(2) = 11$. Therefore, $x_1 = 2 - \frac{6}{11} \approx 1.45$

Newton's Root Finder

Explain the basic idea of Newton's method for finding roots of a function $f(x)$ and show the first step of searching for the root of $f(x) = x^3 - x = 0$, starting from $x_0 = 2$

- The method is based on computing each next approximation x_{n+1} of the goal root using the first-degree Taylor series expansion of $f(x)$ around the current point x_n:

$$f(x) \approx \widehat{f}_1(x; x_n) = f(x_n) + (x - x_n) f'(x_n)$$

- Next approximation x_{n+1} follows from $\widehat{f}_1(x_{n+1}; x_n) = f(x_n) + (x_{n+1} - x_n) f'(x_n) = 0$:

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

- For $f(x) = x^3 - x = 0$, $f'(x) = 3x^2 - 1$; $f(2) = 6$, and $f'(2) = 11$. Therefore, $x_1 = 2 - \frac{6}{11} \approx 1.45$

Solving Linear Systems

Explain in brief, which linear systems $\mathbf{Ax} = \mathbf{b}$ with n unknowns $\mathbf{x}$ and non-singular $n \times n$ matrices $\mathbf{A}$ are easily solvable?

- Systems with diagonal, triangular, and orthonoirmal (or orthogonal) matrices

1. To invert a diagonal matrix, its diagonal elements σ_i are replaced with $\frac{1}{\sigma_i}$; $i = 1, \ldots, n$
2. The system with a triangular matrix is solved sequentially without explicit inversion of $\mathbf{A}$
3. An orthonormal matrix is inverted by transposition: $\mathbf{A}^{-1} = \mathbf{A}^{\mathsf{T}}$

Solving Linear Systems

Explain in brief, which linear systems $\mathbf{Ax} = \mathbf{b}$ with n unknowns $\mathbf{x}$ and non-singular $n \times n$ matrices $\mathbf{A}$ are easily solvable?

- Systems with diagonal, triangular, and orthonoirmal (or orthogonal) matrices

1. To invert a diagonal matrix, its diagonal elements σ_i are replaced with $\frac{1}{\sigma_i}$; $i = 1, \ldots, n$
2. The system with a triangular matrix is solved sequentially without explicit inversion of $\mathbf{A}$
3. An orthonormal matrix is inverted by transposition: $\mathbf{A}^{-1} = \mathbf{A}^{\mathsf{T}}$

Solving Linear Systems

Explain in brief, which linear systems $\mathbf{Ax} = \mathbf{b}$ with n unknowns $\mathbf{x}$ and non-singular $n \times n$ matrices $\mathbf{A}$ are easily solvable?

- Systems with diagonal, triangular, and orthonoirmal (or orthogonal) matrices

1. To invert a diagonal matrix, its diagonal elements σ_i are replaced with $\frac{1}{\sigma_i}$; $i = 1, \ldots, n$
2. The system with a triangular matrix is solved sequentially without explicit inversion of $\mathbf{A}$
3. An orthonormal matrix is inverted by transposition: $\mathbf{A}^{-1} = \mathbf{A}^\mathsf{T}$

Solving Linear Systems

List three essential factorisations of matrices and specify their properties in brief

- Elimination, or LU decomposition: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of lower triangular and upper triangular matrices: $\mathbf{A} = \mathbf{LU}$
- Orthogonalisation, or QR decompposituon: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of orthonormal and upper triangular matrices: $\mathbf{A} = \mathbf{QR}$
- Singular value decomposition: an $m \times n$ matrix $\mathbf{A} = \mathbf{UDV}^\mathsf{T}$ where $\mathbf{U}$ is an $m \times n$ column-orthonormal matrix (left singular vectors), $\mathbf{D}$ is $n \times n$ diagonal matrix of singular values, and $\mathbf{V}$ is an $n \times n$ column-orthonormal matrix (right singular vectors)

Solving Linear Systems

List three essential factorisations of matrices and specify their properties in brief

- Elimination, or LU decomposition: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of lower triangular and upper triangular matrices: $\mathbf{A} = \mathbf{LU}$
- Orthogonalisation, or QR decompostion: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of orthonormal and upper triangular matrices: $\mathbf{A} = \mathbf{QR}$
- Singular value decomposition: an $m \times n$ matrix $\mathbf{A} = \mathbf{UDV}^\mathsf{T}$ where $\mathbf{U}$ is an $m \times n$ column-orthonormal matrix (left singular vectors), $\mathbf{D}$ is $n \times n$ diagonal matrix of singular values, and $\mathbf{V}$ is an $n \times n$ column-orthonormal matrix (right singular vectors)

Solving Linear Systems

List three essential factorisations of matrices and specify their properties in brief

- Elimination, or LU decomposition: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of lower triangular and upper triangular matrices: $\mathbf{A} = \mathbf{LU}$
- Orthogonalisation, or QR decompposituon: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of orthonormal and upper triangular matrices: $\mathbf{A} = \mathbf{QR}$
- Singular value decomposition: an $m \times n$ matrix $\mathbf{A} = \mathbf{UDV}^\mathsf{T}$ where $\mathbf{U}$ is an $m \times n$ column-orthonormal matrix (left singular vectors), $\mathbf{D}$ is $n \times n$ diagonal matrix of singular values, and $\mathbf{V}$ is an $n \times n$ column-orthonormal matrix (right singular vectors)

Solving Linear Systems

List three essential factorisations of matrices and specify their properties in brief

- Elimination, or LU decomposition: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of lower triangular and upper triangular matrices: $\mathbf{A} = \mathbf{LU}$
- Orthogonalisation, or QR decompposituon: an $n \times n$ matrix $\mathbf{A}$ is represented by a product of orthonormal and upper triangular matrices: $\mathbf{A} = \mathbf{QR}$
- Singular value decomposition: an $m \times n$ matrix $\mathbf{A} = \mathbf{UDV}^\mathsf{T}$ where $\mathbf{U}$ is an $m \times n$ column-orthonormal matrix (left singular vectors), $\mathbf{D}$ is $n \times n$ diagonal matrix of singular values, and $\mathbf{V}$ is an $n \times n$ column-orthonormal matrix (right singular vectors)

Constrained Optimisation

How to minimise $f(x, y) = x^4 - 2x^3y + y^4$ under the constraint $x^2 + 8xy + 16y^2 - 25 = 0$? You need only to explain the method and give main relationships, but you need not solve the latter.

- Forming the Lagrangian:

$$\Phi(x, y, \lambda) = x^4 - 2x^3y + y^4 - \lambda(x^2 + 8xy + 16y^2 - 25)$$

- Specify its stationary point:

$$\begin{aligned} \frac{\partial \Phi(x,y,\lambda)}{\partial x} &= 4x^3 - 6x^2y - \lambda(2x + 8y) = 0 \\ \frac{\partial \Phi(x,y,\lambda)}{\partial x} &= -2x^3 + 4y^3 - \lambda(8x + 32y) = 0 \\ & \quad x^2 + 8xy + 16y^2 - 25 = 0 \end{aligned}$$

- Solve the above system for the goal values of λ, x, and y

Constrained Optimisation

How to minimise $f(x,y) = x^4 - 2x^3y + y^4$ under the constraint $x^2 + 8xy + 16y^2 - 25 = 0$? You need only to explain the method and give main relationships, but you need not solve the latter.

- Forming the Lagrangian:

$$\Phi(x,y,\lambda) = x^4 - 2x^3y + y^4 - \lambda(x^2 + 8xy + 16y^2 - 25)$$

- Specify its stationary point:

$$\begin{aligned} \frac{\partial\Phi(x,y,\lambda)}{\partial x} &= 4x^3 - 6x^2y - \lambda(2x + 8y) = 0 \\ \frac{\partial\Phi(x,y,\lambda)}{\partial x} &= -2x^3 + 4y^3 - \lambda(8x + 32y) = 0 \\ & \quad x^2 + 8xy + 16y^2 - 25 = 0 \end{aligned}$$

- Solve the above system for the goal values of λ, x, and y

Least Squares

You have an overdetermined linear system $\mathbf{Ax} = \mathbf{b}$. with n unknowns $\mathbf{Px} = [x_1, \ldots, x_n]^\top$ and $m \gg n$ equations. What is the least squares solution of the system? Derive and explain in brief.

- The least squares solution $\mathbf{x}^*$ minimises the squared error $E(\mathbf{x}) = (\mathbf{b} - \mathbf{Ax})^\top(\mathbf{b} - \mathbf{Ax}) = \mathbf{b}^\top\mathbf{b} - 2\mathbf{x}^\top\mathbf{A}^\top\mathbf{b} + \mathbf{x}^\top\mathbf{A}^\top\mathbf{Ax}$
- This minimiser corresponds to the root of $\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} = 0$, i.e. $\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} = -2\mathbf{A}^\top\mathbf{b} + 2\mathbf{A}^\top\mathbf{Ax} = 0$
- Therefore, $\mathbf{A}^\top\mathbf{Ax} = \mathbf{A}^\top\mathbf{b}$ (the normal equation), and $\mathbf{x}^* = \left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top\mathbf{b}$ if the matrix $\mathbf{A}^\top\mathbf{A}$ is invertible
- Otherwise, the SVD of the $\mathbf{A}^\top\mathbf{A}$ leads to the pseudoinverse matrix to solve the normal equation

Least Squares

You have an overdetermined linear system $\mathbf{Ax} = \mathbf{b}$. with n unknowns $\mathbf{Px} = [x_1, \ldots, x_n]^\mathsf{T}$ and $m \gg n$ equations. What is the least squares solution of the system? Derive and explain in brief.

- The least squares solution $\mathbf{x}^*$ minimises the squared error $E(\mathbf{x}) = (\mathbf{b} - \mathbf{Ax})^\mathsf{T}(\mathbf{b} - \mathbf{Ax}) = \mathbf{b}^\mathsf{T}\mathbf{b} - 2\mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{b} + \mathbf{x}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{Ax}$
- This minimiser corresponds to the root of $\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} = 0$, i.e. $\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} = -2\mathbf{A}^\mathsf{T}\mathbf{b} + 2\mathbf{A}^\mathsf{T}\mathbf{Ax} = 0$
- Therefore, $\mathbf{A}^\mathsf{T}\mathbf{Ax} = \mathbf{A}^\mathsf{T}\mathbf{b}$ (the normal equation), and $\mathbf{x}^* = \left(\mathbf{A}^\mathsf{T}\mathbf{A}\right)^{-1}\mathbf{A}^\mathsf{T}\mathbf{b}$ if the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ is invertible
- Otherwise, the SVD of the $\mathbf{A}^\mathsf{T}\mathbf{A}$ leads to the pseudoinverse matrix to solve the normal equation

Probabilistic Modelling

Given an alphabet $A = \{0, 1, 2\}$, what do you need to define a homogeneous first-order Markov chain of characters from A? Give an example.

- Unconditional probabilities $p(x_0)$, $x_0 \in A$, of the first character, e.g. $p(0) = p(1) = p(2) = \frac{1}{3}$
- Transition probabilities $p(x_i|x_{i-1})$ for the characters at positions $i = 1, 2, \ldots$; $x_i \in A$; $x_{i-1} \in A$, e.g.:

	x_{i-1}		
x_i	0	1	2
0	0.5	0.3	0.7
1	0.1	0.6	0.1
2	0.4	0.1	0.2

Probabilistic Modelling

Given an alphabet $A = \{0, 1, 2\}$, what do you need to define a homogeneous first-order Markov chain of characters from A? Give an example.

- Unconditional probabilities $p(x_0)$, $x_0 \in A$, of the first character, e.g. $p(0) = p(1) = p(2) = \frac{1}{3}$
- Transition probabilities $p(x_i|x_{i-1})$ for the characters at positions $i = 1, 2, \ldots$; $x_i \in A$; $x_{i-1} \in A$, e.g.:

	x_{i-1}		
x_i	0	1	2
0	0.5	0.3	0.7
1	0.1	0.6	0.1
2	0.4	0.1	0.2

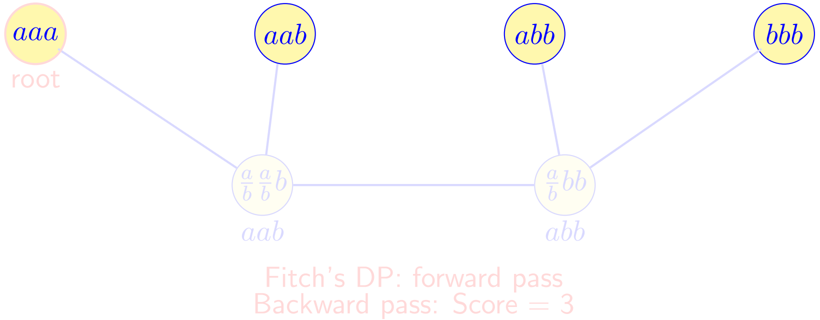
Maximum Parsimony

Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score

Fitch's DP: forward pass
Backward pass: Score = 3

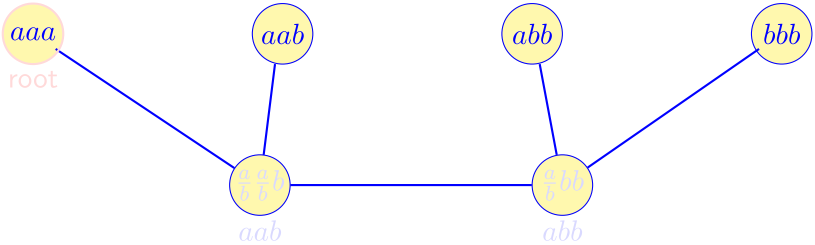
Maximum Parsimony

Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score



Maximum Parsimony

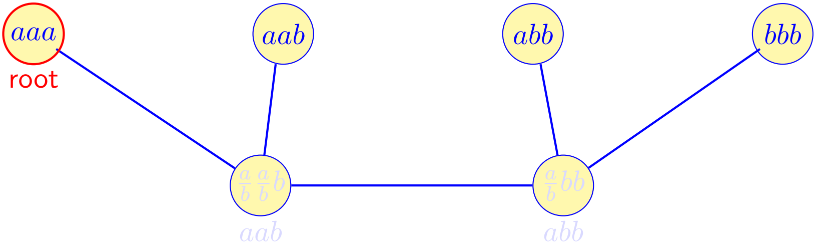
Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score



Fitch's DP: forward pass
Backward pass: Score = 3

Maximum Parsimony

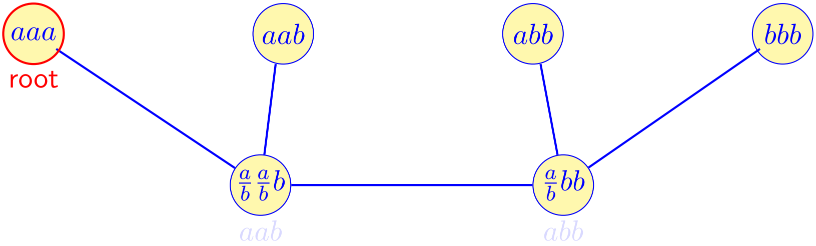
Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score



Fitch's DP: forward pass
Backward pass: Score = 3

Maximum Parsimony

Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score

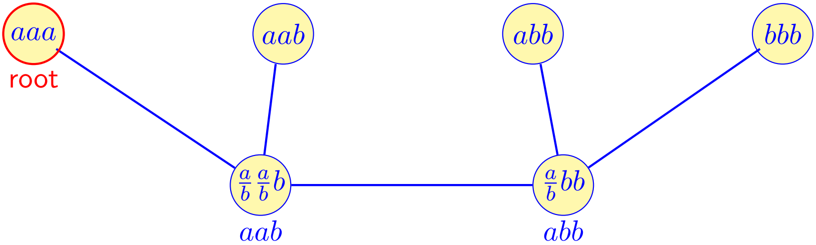


Fitch's DP: forward pass

Backward pass: Score = 3

Maximum Parsimony

Suppose a binary alphabet $A = \{a, b\}$ and 4 strings of length 3: $C = \{aaa, aab, abb, bbb\}$. Draw a Steiner tree with 2 internal nodes and 4 leaves from C such that labels of the internal nodes make the tree with the minimum parsimony score



Fitch's DP: forward pass
Backward pass: Score $= 3$