Simulation and High-Performance Computing

Part 7: Iterative Methods

Steffen Börm

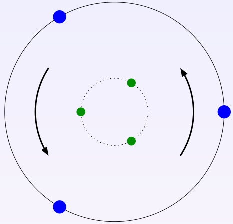
Christian-Albrechts-Universität zu Kiel

October 1st, 2020

Example: Lagrange points

Setting: We have bodies with masses $m_1, \ldots, m_n$ in points $x_1, \ldots, x_n \in \mathbb{R}^2$ in a coordinate system rotating at an angular velocity of α.

Lagrange points: There are points $z \in \mathbb{R}^2$ where the gravitational forces and the centrifugal force cancel each other out.



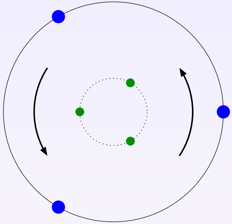
Stable Lagrange points tend to accumulate trojan asteroids of interest to astronomers.

Example: Lagrange points

Setting: We have bodies with masses $m_1, \ldots, m_n$ in points $x_1, \ldots, x_n \in \mathbb{R}^2$ in a coordinate system rotating at an angular velocity of α.

Lagrange points: There are points $z \in \mathbb{R}^2$ where the gravitational forces and the centrifugal force cancel each other out.

Stable Lagrange points tend to accumulate trojan asteroids of interest to astronomers.



Equilibrium of forces described by

$$\alpha^2 z + \gamma \sum_{i=1}^{n} m_i \frac{x_i - z}{\|x_i - z\|^3} = 0.$$

Nonlinear equations like this are generally solved by numerical algorithms.

Iterations

Nonlinear systems can always be written in the form

$$f(x) = 0$$

with a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$.

Problem: Frequently impossible to compute exact solution.

Iterations

Nonlinear systems can always be written in the form

$$f(x) = 0$$

with a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$.

Problem: Frequently impossible to compute exact solution.

Idea: Starting with an initial guess $x_0 \in \mathbb{R}^n$, improve it by applying an iteration function $\Phi\colon \mathbb{R}^n \to \mathbb{R}^n$.

$$x_{m+1} := \Phi(x_m).$$

This procedure may never yield the exact solution, but it can provide us with arbitrarily accurate approximations.

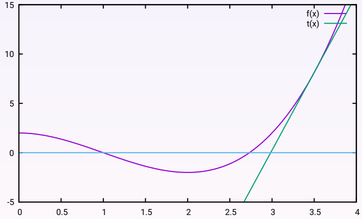
Newton iteration

Idea: Approximate f by its tangent

$$f(x) \approx f(x_m) + f'(x_m)(x - x_m),$$

look for the tangent's zero.

$$0 = f(x_m) + f'(x_m)(x_{m+1} - x_m) \qquad \iff \qquad x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}$$



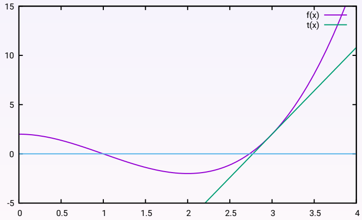
Newton iteration

Idea: Approximate f by its tangent

$$f(x) \approx f(x_m) + f'(x_m)(x - x_m),$$

look for the tangent's zero.

$$0 = f(x_m) + f'(x_m)(x_{m+1} - x_m) \qquad \iff \qquad x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}$$



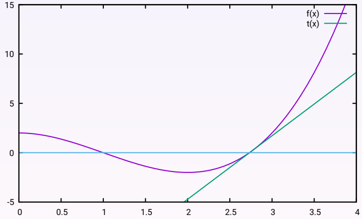
Newton iteration

Idea: Approximate f by its tangent

$$f(x) \approx f(x_m) + f'(x_m)(x - x_m),$$

look for the tangent's zero.

$$0 = f(x_m) + f'(x_m)(x_{m+1} - x_m) \qquad \Longleftrightarrow \qquad x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}$$



Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$0 = f(x) = f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2,$$

Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$0 = f(x) = f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2,$$

$$0 = \frac{f(x_m)}{f'(x_m)} + x - x_m + \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2,$$

Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$
\begin{aligned}
0 = f(x) &= f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2,\\
0 &= \frac{f(x_m)}{f'(x_m)} + x - x_m + \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2,\\
x_{m+1} - x &= \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2.
\end{aligned}
$$

Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$0 = f(x) = f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2,$$
$$0 = \frac{f(x_m)}{f'(x_m)} + x - x_m + \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2,$$
$$x_{m+1} - x = \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2.$$

Convergence: We assume $\frac{|f''(y)|}{2\,|f'(z)|} \leq c$ and obtain $|x_{m+1} - x| \leq c\,|x_m - x|^2$.

Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$
\begin{aligned}
0 = f(x) &= f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2,\\
0 &= \frac{f(x_m)}{f'(x_m)} + x - x_m + \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2,\\
x_{m+1} - x &= \frac{f''(\eta)}{2\,f'(x_m)}(x - x_m)^2.
\end{aligned}
$$

Convergence: We assume $\frac{|f''(y)|}{2\,|f'(z)|} \leq c$ and obtain $|x_{m+1} - x| \leq c\,|x_m - x|^2$.

Example: If $c\,|x_0 - x| \leq \frac{1}{2}$, we find

$$
\begin{aligned}
&c\,|x_1 - x| \leq c^2|x_0 - x|^2 \leq \tfrac{1}{4}, && c\,|x_2 - x| \leq c^2|x_1 - x|^2 \leq \tfrac{1}{16},\\
&c\,|x_3 - x| \leq c^2|x_2 - x|^2 \leq \tfrac{1}{256},
\end{aligned}
$$

Quadratic convergence

Taylor expansion around x_m yields $\eta \in \mathbb{R}$ with

$$\begin{aligned}
0 = f(x) &= f(x_m) + f'(x_m)(x - x_m) + \frac{f''(\eta)}{2}(x - x_m)^2, \\
0 &= \frac{f(x_m)}{f'(x_m)} + x - x_m + \frac{f''(\eta)}{2\, f'(x_m)}(x - x_m)^2, \\
x_{m+1} - x &= \frac{f''(\eta)}{2\, f'(x_m)}(x - x_m)^2.
\end{aligned}$$

Convergence: We assume $\frac{|f''(y)|}{2\,|f'(z)|} \leq c$ and obtain $|x_{m+1} - x| \leq c\,|x_m - x|^2$.

Example: If $c\,|x_0 - x| \leq \frac{1}{2}$, we find

$$\begin{aligned}
c\,|x_1 - x| &\leq c^2 |x_0 - x|^2 \leq \tfrac{1}{4}, & c\,|x_2 - x| &\leq c^2 |x_1 - x|^2 \leq \tfrac{1}{16}, \\
c\,|x_3 - x| &\leq c^2 |x_2 - x|^2 \leq \tfrac{1}{256}, & c\,|x_m - x| &\leq 2^{-2^m}.
\end{aligned}$$

Experiment: Square root

Approach: Square root of $a \in \mathbb{R}_{>0}$ is a zero of $f(x) = x^2 - a$.

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)} = x_m - \frac{x_m^2 - a}{2x_m} = \frac{1}{2}\left(x_m + \frac{a}{x_m}\right).$$

m	$x_0 = 2$	$x_0 = 10$	$x_0 = -5$
1	8.58_{-2}	3.69_{+0}	1.29_{+0}
2	2.45_{-3}	1.33_{+0}	3.06_{-1}
3	2.12_{-6}	3.23_{-1}	2.72_{-2}
4	1.59_{-12}	3.00_{-2}	2.57_{-4}
5		3.12_{-4}	2.34_{-8}
6		3.44_{-8}	2.22_{-16}
7		2.22_{-16}	

Observation: Convergence quadratic once accurate enough.

Experiment: Reciprocal square root

Approach: Reciprocal square root of $a \in \mathbb{R}_{>0}$ is a zero of $f(x) = \frac{1}{x^2} - a$.

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)} = x_m - \frac{x_m^{-2} - a}{-2\,x_m^{-3}} = \frac{x_m}{2}\left(3 - a\,x_m^2\right).$$

m	$x_0 = 1$	$x_0 = 1.5$	$x_0 = 2$
1	2.07_{-1}	4.18_{-1}	4.29_{+0}
2	8.21_{-2}	4.43_{-1}	1.18_{+2}
3	1.37_{-2}	3.30_{-1}	1.62_{+6}
4	3.98_{-4}	1.95_{-1}	4.27_{+18}
5	3.36_{-7}	7.32_{-2}	7.77_{+55}
6	2.40_{-13}	1.10_{-2}	4.70_{+167}
7	1.11_{-16}	2.55_{-4}	∞
8	0	1.38_{-7}	

Observation: Convergence quadratic once accurate enough, no convergence for unsuitable initial guesses x_0.

Multidimensional Newton

Idea: Replace the tangent by tangential hyperplanes.

$$0 = f(x) \approx f(x_m) + Df(x_m)(x - x_m)$$

with the Jacobian matrix

$$Df(z) = \begin{pmatrix} \frac{\partial f_1}{\partial z_1}(z) & \cdots & \frac{\partial f_1}{\partial z_n}(z) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial z_1}(z) & \cdots & \frac{\partial f_n}{\partial z_n}(z) \end{pmatrix}.$$

Result: Solving the approximated problem yields

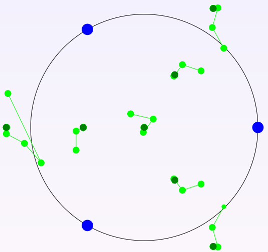
$$x_{m+1} = x_m - Df(x_m)^{-1} f(x_m).$$

We have to solve a linear system of equations in each step.

Experiment: Lagrange points

Goal: Find solutions of

$$\alpha^2 x + \gamma \sum_{i=1}^{n} \frac{y_i - x}{\|y_i - x\|^3} = 0.$$



Observation: Different starting points lead to different results.

Example: Resonance frequencies

Goal: Find non-trivial solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c\Delta u(t,x).$$

Example: Resonance frequencies

Goal: Find non-trivial solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c\Delta u(t,x).$$

Approach: Separate variables $u(t,x) = \cos(\omega t)\, e(x)$.

$$\omega^2 e(x) = -c\Delta e(x).$$

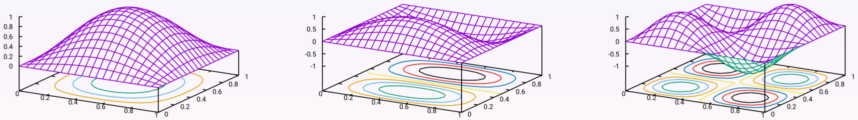
Example: Resonance frequencies

Goal: Find non-trivial solution of the wave equation

$$\frac{\partial^2 u}{\partial t^2}(t,x) = c\Delta u(t,x).$$

Approach: Separate variables $u(t,x) = \cos(\omega t)\, e(x)$.

$$\omega^2 e(x) = -c\Delta e(x).$$



Eigenvalue problems

Task: Given a matrix $A \in \mathbb{R}^{n \times n}$, find $\lambda \in \mathbb{R}$ and $e \in \mathbb{R}^n$ with

$$Ae = \lambda\, e, \qquad\qquad e \neq 0$$

λ is called an eigenvalue, e an eigenvector.

Variations:

- Find the smallest or largest eigenvalue and a matching eigenvector.
- Find an entire basis of eigenvectors.
- Find a basis for an invariant subspace.

Power iteration

Approach: If a vector $x \in \mathbb{R}^n$ can be written as

$$x = e_1 + e_2 + \ldots + e_k$$

with eigenvectors $e_1, \ldots, e_k$ for eigenvalues $\lambda_1, \ldots, \lambda_k$, we have

$$\begin{aligned} A^m x &= A^m e_1 + A^m e_2 + \ldots + A^m e_k \\ &= \lambda_1^m e_1 + \lambda_2^m e_2 + \ldots + \lambda_k^m e_k. \end{aligned}$$

Idea: If $|\lambda_1| > |\lambda_2|, \ldots, |\lambda_k|$, the first component in the sum will dominate if m is large enough, i.e., we will have an approximation of e_1.

Power iteration

Approach: If a vector $x \in \mathbb{R}^n$ can be written as

$$x = e_1 + e_2 + \ldots + e_k$$

with eigenvectors $e_1, \ldots, e_k$ for eigenvalues $\lambda_1, \ldots, \lambda_k$, we have

$$\begin{aligned} A^m x &= A^m e_1 + A^m e_2 + \ldots + A^m e_k \\ &= \lambda_1^m e_1 + \lambda_2^m e_2 + \ldots + \lambda_k^m e_k. \end{aligned}$$

Idea: If $|\lambda_1| > |\lambda_2|, \ldots, |\lambda_k|$, the first component in the sum will dominate if m is large enough, i.e., we will have an approximation of e_1.

Practice: To avoid an overflow, compute the normalized sequence

$$\hat{x}_{m+1} := A x_m, \qquad x_{m+1} := \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\tan^2 \angle(x_m, e) = \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2}$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\begin{aligned}
\tan^2 \angle(x_m, e) &= \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}\|e_2\|^2 + \ldots + |\lambda_k|^{2m}\|e_k\|^2}{|\lambda_1|^{2m}\|e_1\|^2}
\end{aligned}$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\begin{aligned}
\tan^2 \angle(x_m, e) &= \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2} \\
&= \frac{|\lambda_2|^{2m} \|e_2\|^2 + \ldots + |\lambda_k|^{2m} \|e_k\|^2}{|\lambda_1|^{2m} \|e_1\|^2} \\
&\leq \frac{|\lambda_2|^{2m} \|e_2\|^2 + \ldots + |\lambda_2|^{2m} \|e_k\|^2}{|\lambda_1|^{2m} \|e_1\|^2}
\end{aligned}$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\begin{aligned}
\tan^2 \angle(x_m, e) &= \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2} \\
&= \frac{|\lambda_2|^{2m} \|e_2\|^2 + \ldots + |\lambda_k|^{2m} \|e_k\|^2}{|\lambda_1|^{2m} \|e_1\|^2} \\
&\leq \frac{|\lambda_2|^{2m} \|e_2\|^2 + \ldots + |\lambda_2|^{2m} \|e_k\|^2}{|\lambda_1|^{2m} \|e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}}{|\lambda_1|^{2m}} \frac{\|e_2\|^2 + \ldots + \|e_k\|^2}{\|e_1\|^2}
\end{aligned}$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\begin{aligned}
\tan^2 \angle(x_m, e) &= \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}\|e_2\|^2 + \ldots + |\lambda_k|^{2m}\|e_k\|^2}{|\lambda_1|^{2m}\|e_1\|^2} \\
&\leq \frac{|\lambda_2|^{2m}\|e_2\|^2 + \ldots + |\lambda_2|^{2m}\|e_k\|^2}{|\lambda_1|^{2m}\|e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}}{|\lambda_1|^{2m}} \frac{\|e_2\|^2 + \ldots + \|e_k\|^2}{\|e_1\|^2} = \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^{2m} \tan^2 \angle(x_0, e).
\end{aligned}$$

Convergence

Approach: If $e_1, \ldots, e_k$ is an orthogonal basis of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ satisfying $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, we can bound the angle between x_m and e_1 by

$$\begin{aligned}
\tan^2 \angle(x_m, e) &= \frac{\|\lambda_2^m e_2 + \ldots + \lambda_k^m e_k\|^2}{\|\lambda_1^m e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}\|e_2\|^2 + \ldots + |\lambda_k|^{2m}\|e_k\|^2}{|\lambda_1|^{2m}\|e_1\|^2} \\
&\leq \frac{|\lambda_2|^{2m}\|e_2\|^2 + \ldots + |\lambda_2|^{2m}\|e_k\|^2}{|\lambda_1|^{2m}\|e_1\|^2} \\
&= \frac{|\lambda_2|^{2m}}{|\lambda_1|^{2m}} \frac{\|e_2\|^2 + \ldots + \|e_k\|^2}{\|e_1\|^2} = \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^{2m} \tan^2 \angle(x_0, e).
\end{aligned}$$

Result: Convergence rate $\frac{|\lambda_2|}{|\lambda_1|}$ for the tangent.

Inverse iteration

Task: Frequently we need not the largest, but the smallest eigenvalue, e.g., when computing the fundamental frequency of a system.

Idea: Work with A^{-1} instead of A.

$$\begin{aligned} A^{-m}x &= A^{-m}e_1 + A^{-m}e_2 + \ldots + A^{-m}e_k \\ &= \tfrac{1}{\lambda_1^m}e_1 + \tfrac{1}{\lambda_2^m}e_2 + \ldots + \tfrac{1}{\lambda_k^m}e_k. \end{aligned}$$

If $|\lambda_1| < |\lambda_2|, \ldots, |\lambda_k|$, the first component in the sum will dominate if m is large enough.

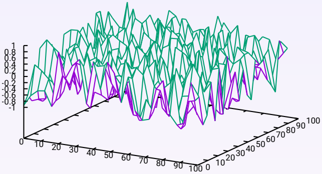
Practice: We have to solve a linear system in each step.

$$A\hat{x}_{m+1} = x_m, \qquad\qquad x_{m+1} := \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

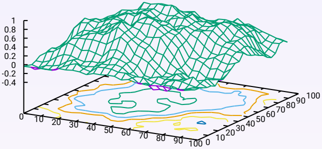
m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	



Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

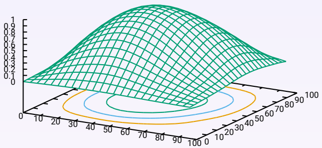
m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	
1	3.87_{-1}	4.19_{-1}	166.3



Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

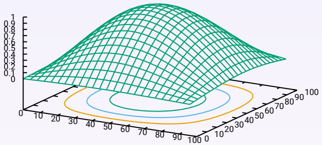
m	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	
1	3.87_{-1}	4.19_{-1}	166.3
2	1.20_{-1}	1.20_{-1}	3.5



Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

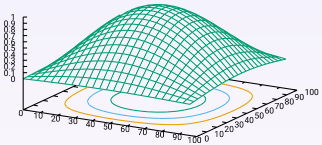
m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	
1	3.87_{-1}	4.19_{-1}	166.3
2	1.20_{-1}	1.20_{-1}	3.5
3	4.46_{-2}	4.46_{-2}	2.7



Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

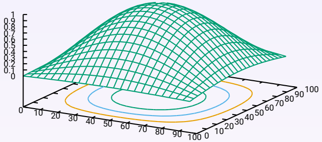
m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	
1	3.87_{-1}	4.19_{-1}	166.3
2	1.20_{-1}	1.20_{-1}	3.5
3	4.46_{-2}	4.46_{-2}	2.7
4	1.73_{-2}	1.73_{-2}	2.6



Experiment: Inverse iteration

Goal: Find the smallest eigenvalue of the wave equation.

m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	6.97_{+1}	
1	3.87_{-1}	4.19_{-1}	166.3
2	1.20_{-1}	1.20_{-1}	3.5
3	4.46_{-2}	4.46_{-2}	2.7
4	1.73_{-2}	1.73_{-2}	2.6
5	6.86_{-3}	6.86_{-3}	2.5
6	2.73_{-3}	2.73_{-3}	2.5
7	1.09_{-3}	1.09_{-3}	2.5
8	4.36_{-4}	4.36_{-4}	2.5



Inverse iteration with shift

Task: Sometimes we need the eigenvalue closest to a number μ, e.g., to find resonances close to the frequency of an incident wave.

Idea: Work with $(A - \mu I)^{-1}$ instead of A.

$$\begin{aligned}(A - \mu I)^{-m} x &= (A - \mu I)^{-m} e_1 + (A - \mu I)^{-m} e_2 + \ldots + (A - \mu I)^{-m} e_k \\ &= \tfrac{1}{(\lambda_1 - \mu)^m} e_1 + \tfrac{1}{(\lambda_2 - \mu)^m} e_2 + \ldots + \tfrac{1}{(\lambda_k - \mu)^m} e_k.\end{aligned}$$

If $|\lambda_1 - \mu| < |\lambda_2 - \mu|, \ldots, |\lambda_k - \mu|$, the first component in the sum will dominate if m is large enough.

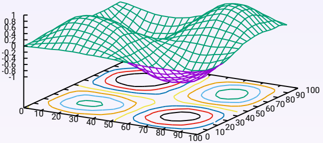
Practice: We have to solve a linear system in each step.

$$(A - \mu I)\hat{x}_{m+1} = x_m, \qquad x_{m+1} := \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Experiment: Inverse iteration with shift

Goal: Find the eigenvalue closest to $\mu = 75$ of the wave equation.

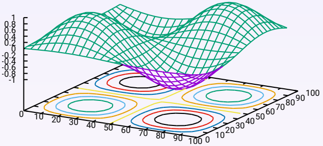
m	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	1.00_{+0}	8.30_{+1}	
1	1.73_{-1}	1.76_{-1}	472.6



Experiment: Inverse iteration with shift

Goal: Find the eigenvalue closest to $\mu = 75$ of the wave equation.

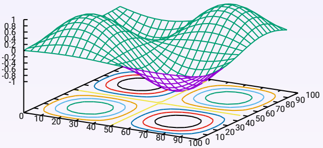
m	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	1.00_{+0}	8.30_{+1}	
1	1.73_{-1}	1.76_{-1}	472.6
2	2.08_{-2}	2.09_{-2}	8.4



Experiment: Inverse iteration with shift

Goal: Find the eigenvalue closest to $\mu = 75$ of the wave equation.

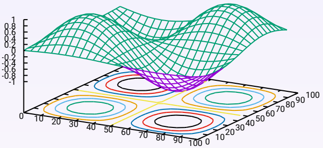
m	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	1.00_{+0}	8.30_{+1}	
1	1.73_{-1}	1.76_{-1}	472.6
2	2.08_{-2}	2.09_{-2}	8.4
3	3.04_{-3}	3.04_{-3}	6.9



Experiment: Inverse iteration with shift

Goal: Find the eigenvalue closest to $\mu = 75$ of the wave equation.

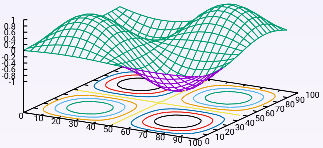
m	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	1.00_{+0}	8.30_{+1}	
1	1.73_{-1}	1.76_{-1}	472.6
2	2.08_{-2}	2.09_{-2}	8.4
3	3.04_{-3}	3.04_{-3}	6.9
4	4.64_{-4}	4.64_{-4}	6.6



Experiment: Inverse iteration with shift

Goal: Find the eigenvalue closest to $\mu = 75$ of the wave equation.

m	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	1.00_{+0}	8.30_{+1}	
1	1.73_{-1}	1.76_{-1}	472.6
2	2.08_{-2}	2.09_{-2}	8.4
3	3.04_{-3}	3.04_{-3}	6.9
4	4.64_{-4}	4.64_{-4}	6.6
5	7.15_{-5}	7.15_{-5}	6.5
6	1.11_{-5}	1.11_{-5}	6.5
7	1.72_{-5}	1.72_{-6}	6.4
8	2.67_{-7}	2.67_{-7}	6.4



Rayleigh quotient

Task: Assuming that we have an approximation x of an eigenvector e, how do we find an approximation of the corresponding eigenvalue λ?

Idea: Use the inner product. For the exact eigenvector e, we have

$$Ae = \lambda e, \qquad \langle e, Ae \rangle = \lambda \langle e, e \rangle, \qquad \frac{\langle e, Ae \rangle}{\langle e, e \rangle} = \lambda.$$

Rayleigh quotient

Task: Assuming that we have an approximation x of an eigenvector e, how do we find an approximation of the corresponding eigenvalue λ?

Idea: Use the inner product. For the exact eigenvector e, we have

$$Ae = \lambda e, \qquad \langle e, Ae \rangle = \lambda \langle e, e \rangle, \qquad \frac{\langle e, Ae \rangle}{\langle e, e \rangle} = \lambda.$$

Rayleigh quotient: For the approximated eigenvector, we find

$$\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \approx \lambda.$$

Rayleigh quotient

Task: Assuming that we have an approximation x of an eigenvector e, how do we find an approximation of the corresponding eigenvalue λ?

Idea: Use the inner product. For the exact eigenvector e, we have

$$Ae = \lambda e, \qquad \langle e, Ae \rangle = \lambda \langle e, e \rangle, \qquad \frac{\langle e, Ae \rangle}{\langle e, e \rangle} = \lambda.$$

Rayleigh quotient: For the approximated eigenvector, we find

$$\frac{\langle x, Ax \rangle}{\langle x, x \rangle} \approx \lambda.$$

Stopping criterion: The Rayleigh quotient can be used to estimate the accuracy of an approximation by checking

$$\|Ax - \tilde{\lambda} x\| \leq \epsilon \, \|x\| \qquad \text{with} \qquad \tilde{\lambda} := \frac{\langle x, Ax \rangle}{\langle x, x \rangle}.$$

Rayleigh iteration

Task: Find good shift parameters μ for the inverse iteration.

Idea: Use the Rayleigh quotient, since it provides approximate eigenvalues.

$$\mu_m := \frac{\langle x_m, A x_m \rangle}{\langle x_m, x_m \rangle}, \qquad (A - \mu_m I)\hat{x}_{m+1} = x_m, \qquad x_{m+1} := \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Rayleigh iteration

Task: Find good shift parameters μ for the inverse iteration.

Idea: Use the Rayleigh quotient, since it provides approximate eigenvalues.

$$\mu_m := \frac{\langle x_m, Ax_m \rangle}{\langle x_m, x_m \rangle}, \qquad (A - \mu_m I)\hat{x}_{m+1} = x_m, \qquad x_{m+1} := \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

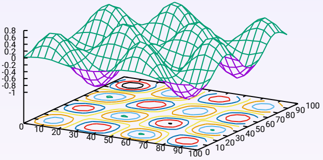
Consequences:

- In each step, a different matrix $A - \mu_m I$ has to be used.
- If the initial vector x_0 is close enough to the desired eigenvector, we obtain quadratic or even cubic convergence.
- If the initial vector x_0 is not close enough to the desired eigenvector, we may observe convergence to another eigenvector or even no convergence at all.

Experiment: Rayleigh iteration

Goal: Find the eigenvalue closest to $\mu = 310$ of the wave equation.

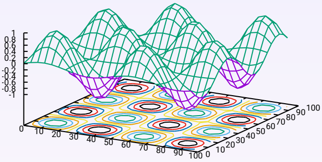
m	μ	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	3.10_{+2}	1.00_{+0}	5.43_{+1}	
1	3.20_{+2}	2.03_{-1}	2.07_{-1}	261.9



Experiment: Rayleigh iteration

Goal: Find the eigenvalue closest to $\mu = 310$ of the wave equation.

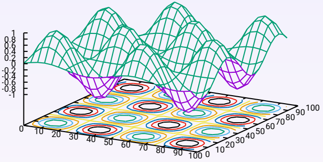
m	μ	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	3.10_{+2}	1.00_{+0}	5.43_{+1}	
1	3.20_{+2}	2.03_{-1}	2.07_{-1}	261.9
2	3.15_{+2}	4.10_{-2}	4.10_{-2}	5.1



Experiment: Rayleigh iteration

Goal: Find the eigenvalue closest to $\mu = 310$ of the wave equation.

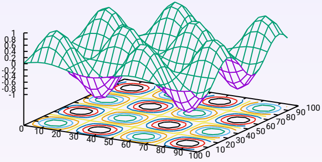
m	μ	$\sin\alpha_m$	$\tan\alpha_m$	ratio
0	3.10_{+2}	1.00_{+0}	5.43_{+1}	
1	3.20_{+2}	2.03_{-1}	2.07_{-1}	261.9
2	3.15_{+2}	4.10_{-2}	4.10_{-2}	5.1
3	3.15_{+2}	3.85_{-5}	3.84_{-5}	1067.6



Experiment: Rayleigh iteration

Goal: Find the eigenvalue closest to $\mu = 310$ of the wave equation.

m	μ	$\sin \alpha_m$	$\tan \alpha_m$	ratio
0	3.10_{+2}	1.00_{+0}	5.43_{+1}	
1	3.20_{+2}	2.03_{-1}	2.07_{-1}	261.9
2	3.15_{+2}	4.10_{-2}	4.10_{-2}	5.1
3	3.15_{+2}	3.85_{-5}	3.84_{-5}	1067.6
4	3.15_{+2}	3.06_{-13}	3.06_{-13}	1.25_{+8}



Summary

Newton iteration can be used to solve nonlinear equations $f(x) = 0$.

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)}, \qquad x_{m+1} = x_m - Df(x_m)^{-1} f(x_m).$$

If x_m is sufficiently close to the solution, we have quadratic convergence.

Power iteration can be used to solve eigenvalue problems $Ae = \lambda e$.

$$\hat{x}_{m+1} = A x_m, \qquad x_{m+1} = \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Inverse iteration can be used to choose specific eigenvalues.

$$(A - \mu I)\hat{x}_{m+1} = x_m, \qquad x_{m+1} = \frac{\hat{x}_{m+1}}{\|\hat{x}_{m+1}\|}.$$

Rayleigh quotient provides an approximation of an eigenvalue.