Notes

- Extra class this Friday 1-2pm
- Assignment 2 is out
- Error in last lecture: quasi-Newton methods based on the secant condition:
  $$H\Delta x \approx g(x+\Delta x) - g(x)$$
  - Really just Taylor series applied to the gradient function g:
  $$g(x+\Delta x) = g(x) + H\Delta x + O(\Delta x^2)$$

Nonlinear Least Squares

- One particular nonlinear problem of interest:
  $$\min_x \|b - g(x)\|_2^2$$
- Can we exploit the structure of this objective?
- Gradient: $-2J^T(b - g(x)), \quad J_{ij} = \frac{\partial g_i}{\partial x_j}$
- Hessian: $H = 2J^T J - 2(b - g(x)) \bullet \frac{\partial^2 g}{\partial x^2}$

Gauss-Newton

- Assuming g is close to linear, just keep the $J^T J$ term
- Automatically positive (semi-)definite: in fact, modified Newton solve is now a linear least-squares problem!
  $$\min_{\Delta x} \|(b - g(x)) - J\Delta x\|_2^2$$
- Convergence may not be quite as fast as Newton, but more robust.

Time Integration

- A core problem across many applications
  - Physics, robotics, finance, control, …
- Typical statement:
  system of ordinary differential equations, first order $\frac{dy}{dt} = f(y,t)$

  Subject to initial conditions:
  $$y(0) = y_0$$
- Higher order equations can be reduced to a system - though not always a good thing to do

Forward Euler

- Derive from Taylor series:
  $$y(t+\Delta t) = y(t) + \Delta t \frac{dy}{dt}(t) + O(\Delta t^2)$$
  $$y_{n+1} = y_n + \Delta t\, f(y_n, t_n)$$
- **Truncation error** is second-order
- (Global) accuracy is first-order
  - Heuristic: to get to a fixed time T, need T/Δt time steps

Analyzing Forward Euler

- Euler proved convergence as Δt goes to 0
- In practice, need to pick Δt > 0
  - how close to zero do you need to go?
- The first and foremost tool is using a simple model problem that we can solve exactly
  $$\frac{dy}{dt} = Ay$$

Solving Linear ODE's

- First: what is the exact solution?
- For scalars: $\frac{dy}{dt} = Ay, \quad y(0) = y_0$

  $\Rightarrow \quad y(t) = e^{At} y_0$
- This is in fact true for systems as well, though needs the "matrix exponential"
- More elementary approach:
  change basis to diagonalize A
  (or reduce to Jordan canonical form)

The Test Equation

- We can thus simplify our model down to a single scalar equation, **the test equation**

  $$\frac{dy}{dt} = \lambda y$$
- Relate this to a full nonlinear system by taking the eigenvalues of the Jacobian of f:

  $$\left(\frac{\partial f}{\partial y}\right) x = \lambda x$$
  - But obviously nonlinearities may cause unexpected things to happen:
    full nonlinear analysis is generally problem-specific, and hard or impossible

Solution of the Test Equation

- Split scalar into real and imaginary parts:

  $$\lambda = a + b\sqrt{-1}$$
- For initial conditions y_0=1

  $$y(t) = e^{\lambda t} = e^{at}\left(\cos bt + \sqrt{-1}\sin bt\right)$$
- The magnitude only depends on a
  - If a<0: exponential decay (stable)
    If a>0: exponential increase (unstable)
    If a=0: borderline-stable