Today

- non linearity
- optimization

Announcements

- HW8
- 4G4
Fixed Point Iteration

\[ \ln |D| \leq \left| \frac{\partial f}{\partial x} \right| < 1 \]

\[ g'(x) \neq 0 \]

When does this converge?

\[ x_0 = \text{starting guess} \]

\[ x_{k+1} = g(x_k) \]

\[ f: \mathbb{R}^n \rightarrow \mathbb{R}^n \]

\[ g: \mathbb{R}^n \rightarrow \mathbb{R}^n \]

\[ \|g\| \leq 1 \Rightarrow g(x^*) \leq 1 \Rightarrow \text{lin. conv.} \]
\[ \frac{\partial g(x^*)}{\partial x} = 0 \ \Rightarrow \ \text{quadratic convergence} \]

\[ \frac{\partial f(x + \hat{h})}{\partial x} = f'(x) + \int g'(x) \cdot \hat{h} \]
\[ f_1 (x_1, x_2) = 0 \]

\[ f_2 (x_1, x_2) = 0 \]
Newton’s Method

What does Newton’s method look like in $n$ dimensions?

\[ f(x_n + h) = f(x_n) + \nabla f(x_n) \cdot h + o(h) \]

- \( f(x_n) - \nabla f(x_n) \cdot h \)
- \( \nabla f(x_n)^{-1} \cdot f(x_n) = h \)

\[ x_{n+1} = x_n - \frac{f'(x_n) f(x_n)}{f'(x_n)^T f(x_n)} \]

Downsides of $n$-dim. Newton?

- local conv.
- \( \nabla f \)

**Demo:** Newton’s method in $n$ dimensions
Secant in $n$ dimensions?

What would the secant method look like in $n$ dimensions?

\[
\begin{align*}
\hat{x}_k &= \frac{f(x_k)}{f(x_k) - f(x_{k-1})}, \\
\text{first column of } J_k
\end{align*}
\]

\[
\tilde{J}_0 = I \quad \text{with eval } f(x_1), f(x_0) \\
\text{update } J_1
\]

Broyden's method
Outline

Introduction to Scientific Computing

Systems of Linear Equations

Linear Least Squares

Eigenvalue Problems

Nonlinear Equations

Optimization
- Methods for unconstrained opt. in one dimension
- Methods for unconstrained opt. in n dimensions
- Nonlinear Least Squares
- Constrained Optimization

Interpolation

Numerical Integration and Differentiation

Initial Value Problems for ODEs

Boundary Value Problems for ODEs

Partial Differential Equations and Sparse Linear Algebra

Fast Fourier Transform
Optimization: Problem Statement

*Have:* Objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$

*Want:* Minimizer $x^* \in \mathbb{R}^n$ so that

$$f(x^*) = \min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \quad g(x) = 0 \quad \text{and} \quad h(x) \leq 0.$$
Optimization: Observations

Q: What if we are looking for a maximizer not a minimizer? Give some examples:

- training ANNs
- robot path planning

What about multiple objectives?

- combine them
- look up Pareto optimal
Existence/Uniqueness

Terminology: global minimum / local minimum

Under what conditions on $f$ can we say something about existence/uniqueness?

If $f : S \rightarrow \mathbb{R}$ is continuous on a closed and bounded set $S \subseteq \mathbb{R}^n$, then

$a \text{ minimum exists.}$

$f : S \rightarrow \mathbb{R}$ is called coercive on $S \subseteq \mathbb{R}^n$ (which must be unbounded) if

$$\lim_{\|x\| \to \infty} f(x) = +\infty$$

If $f$ is coercive, .......

$a \text{ global minimum.}$
Convexity

$S \subseteq \mathbb{R}^n$ is called convex if for all $x, y \in S$ and all $0 \leq \alpha \leq 1$

\[
\alpha x + (1-\alpha)y \in S.
\]

$f : S \to \mathbb{R}$ is called convex on $S \subseteq \mathbb{R}^n$ if for \( x, y \in S \) and all $0 \leq \alpha \leq 1$

\[
f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)
\]

Q: Give an example of a convex, but not strictly convex function.

\[\rightarrow\] likely non-uniqueness of the min.
Convexity: Consequences

If $f$ is convex, ...

- $f$ is continuous
- any local min is a global min.

If $f$ is strictly convex, ...

any local min is a unique global min.
Optimality Conditions

If we have found a candidate $x^*$ for a minimum, how do we know it actually is one? Assume $f$ is smooth, i.e. has all needed derivatives.

In one-d.

Necessary: $f'(x^*) = 0$

Sufficient: $f''(x^*) > 0$ implies local min.

In multi-d.,

Necessary: $\nabla f = 0$

Sufficient: $\nabla f = 0$ and $H_f(x^*)$ pos. def.
\[ \partial_{x_j x_i} \mathcal{F} \]

\[ H_\varphi = \begin{pmatrix} 0_{x_1 x_2} & \partial_{x_2 x_1} \mathcal{F} \\ \partial_{x_1 x_2} \mathcal{F} & 0_{x_1 x_2} \end{pmatrix} \]
Optimization: Observations

Q: Come up with a hypothetical approach for finding minima.

\[
\text{Solve } \nabla f(x) = 0 \quad \text{e.g. using Newton}
\]

Q: Is the Hessian symmetric?

yes

Q: How can we practically test for positive definiteness?

Cholesky.

\[x^T A x > 0\]
Sensitivity and Conditioning (1D)

How does optimization react to a slight perturbation of the minimum?

\[ \rho (x^* + h) = \rho (x^*) + \rho'(x^*) h + \rho''(x^*) \frac{h^2}{2} + O(h^3) \]

\[ |x - x^*| \leq \sqrt{2 \varepsilon / \rho''(x^*)} \]

\[ \rho \varepsilon = 10^{-16} \implies 20 \text{ on solution is} 10^{-8} \]
Sensitivity and Conditioning (nD)

How does optimization react to a slight perturbation of the minimum?