Recap

- Homotopy and Bifurcation
Recap

\[
f(x) = \begin{pmatrix} (x_1 - 1)^2 + x_2^2 - 1 \\ (x_1 + 1)^2 + x_2^2 - 1 \end{pmatrix} = 0
\]

\[
J(x) = \begin{pmatrix} 2(x_1 - 1) & 2x_2 \\ 2(x_1 + 1) & 2x_2 \end{pmatrix}
\]
Optimization

• Problems of the sort:

\[
\min_{x \in D} f(x) \quad \text{arg min}_{x \in D} f(x)
\]

• \( f(x) \): objective function, cost function, energy
  • “metric to compare alternatives”

• \( x \): “design alternatives”

• \( D \): feasible set

• Maximization of \( f(x) \) is just minimization of \( -f(x) \)
Optimization

\[ f(x_1, x_2) \]
Optimization

• Goal: find \( x^* \in D : f(x^*) < f(x) \quad \forall x \in D \)

• \( x^* \) is not necessarily unique. There could be more than one \( x^* \) in \( D \).

• Convexity: a function is convex if the line connecting any two points above the function is also above the function:

\[
\begin{align*}
x^* &\in D : f(x^*) < f(x) \quad \forall x \in D \\
\text{convex} &\quad \text{non-convex}
\end{align*}
\]

• Convex functions have a single, global minimum

• Most algorithms are characterized in terms of their ability to find the global minimum of convex functions.

• Non-convex function may have global or local minima
Optimization

• Examples:
  • Find the value of $x$ that minimizes
    \[ f(x) = x^2 + 2x + 1 \]
  • Find the value of $x \in [0, 1]$ that minimizes
    \[ f(x) = x^2 + 2x + 1 \]
Optimization

• Examples: linear programs
  - Premium and regular ice cream are sold for $5/gallon and $3.5/gallon respectively.
  - Premium ice cream is 30% air by volume while regular ice cream is 50% air by volume.
  - We can produce $X$ gallons of premium and $Y$ gallons of regular ice cream all at the same cost, $1/gallon.
  - What fraction of milk processed should go toward premium versus regular ice cream?
Optimization

• $x^* \in D$ is a local minimum of
  • if $\exists \epsilon > 0 : f(x^*) < f(x), \quad \forall x \in D \cap B_\epsilon(x^*)$

• Global minima are also local minima

• If $f(x)$ is convex in $D$ then a local minimum is the global minimum in $D$.

• If $D$ is a closed set, the problem of finding the minimum is called constrained optimization.

• If $D$ is an open set: $\mathbb{R}^N$, the problem of finding the minimum is called unconstrained optimization
Unconstrained Optimization

- Optimality criteria:
  - How do I check for local minima?
  - Assume \( f(x) \) is twice differentiable, then:

\[
f(x + d) = f(x) + g(x)^T d + \frac{1}{2} d^T H(x) d + \ldots
\]

- where: \( g_i(x) = \frac{\partial f}{\partial x_i} \) \( H_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j} \)
- As \( \|d\|_p \to 0 \)

\[
f(x + d) - f(x) = g^T d
\]
- If \( g^T d > 0 \), then \( f(x + d) > f(x) \)
- But, replace \( d \) with \( -d \), and the converse is true
- Therefore, I have a critical point when: \( g = \nabla f(x) = 0 \)
Unconstrained Optimization

- Solving unconstrained optimization problems is the same as solving the system of nonlinear equations:
  \[ g = \nabla f(x) = 0 \]

- Except, we want to ensure that we only find the roots associated with local minima in \( f(x) \)

- If the eigenvalues of the Hessian are positive, we can be sure that \( f(x) \) is a minimum. Why?

- For a minimum, the eigenvalues must be non-negative

- How do we craft an algorithm that only finds minima?
Unconstrained Optimization

- Examples:
  - Calculate the gradient. Where is the critical point?
  - Calculate the Hessian. Is the critical point a minimum?

\[ f(x) = x_1^2 + x_2^2 \]

\[ f(x) = x_1^2 - x_2^2 \]

\[ f(x) = x_1^4 + x_2^4 \]
Unconstrained Optimization

• Method of steepest decent:
  • Solve the equation: $g(x) = \nabla f(x) = 0$, iteratively by taking steps in a direction that decreases $f(x)$

$$x_{i+1} = x_i + \alpha_i d_i$$

• with $\alpha_i > 0$ and $g(x_i)^T d_i < 0$

• This ensures that $d_i$ is a descent direction:

$$f(x_i + \alpha_i d_i) = f(x_i) + \alpha_i g(x_i)^T d_i + \ldots$$

• Which descent direction should I choose?
  • One option: maximize $-g(x_i)^T d_i$
  • C-S inequality: $-g(x_i)^T d_i \leq \|g(x_i)\|_2 \|d_i\|_2$
  • Solution: let $d_i = -g(x_i)$
Unconstrained Optimization

- Method of steepest decent:
  - Example: \( f(x) = x_1^2 + x_2^2 \)
  - Contours for the function:

\[
x_{i+1} = x_i - \alpha_i g(x_i)
\]

Is there a best value of \( \alpha_i \) to use with this function?
Unconstrained Optimization

• Method of steepest decent:
  • Direction of steepest descent: \( d_i = -g(x_i) \)
  • Iterative solution: \( x_{i+1} = x_i - \alpha_i g(x_i) \)
  • For small, positive values of \( \alpha_i \), the iterates continue to reduce \( f(x) \) until \( g(x) = 0 \)
  • The iterative method converges to local minima and potentially saddle points. Need to check the Hessian still to be sure of minima.

• How do I choose values for \( \alpha_i \)?
  • Ideally, we pick the \( \alpha_i \) that leads to the smallest value of \( f(x_{i+1}) \), but this is its own optimization.
  • We can approximate the solution with a line search like in damped Newton-Raphson.
Unconstrained Optimization

- Method of steepest decent:
  - Example: $f(x) = x_1^2 + 10x_2^2$
  - Contours for the function:

Draw the path given by small $\alpha_i$

- The choice of $\alpha_i$ is critical!
- Too small and the convergence is slow
Unconstrained Optimization

- Method of steepest decent:
  - Example: \( f(x) = x_1^2 + 10x_2^2 \)
  - Contours for the function:

\[
\begin{align*}
\text{Draw the path given by larger } & \alpha_i \\
\text{The choice of } & \alpha_i \text{ is critical!} \\
\text{Too big and convergence is erratic}
\end{align*}
\]
Unconstrained Optimization

- Method of steepest decent:
  - Example: $f(x) = x_1^2 + 10x_2^2$
  - Contours for the function:

\[ x_{i+1} = x_i - \alpha_i \nabla f(x_i) \]
Unconstrained Optimization

- Method of steepest descent:
  - Estimating an optimal $\alpha_i$:
    \[
    x_{i+1} = x_i - \alpha_i g(x_i)
    \]
  - Use a Taylor expansion:
    \[
    f(x_{i+1}) = f(x_i) - \alpha_i g(x_i)^T g(x_i) + \frac{1}{2} \alpha_i^2 g(x_i)^T H(x_i) g(x_i) + \ldots
    \]
    - This is quadratic in $\alpha_i$, so find the minimum:
      \[
      \alpha_i = \frac{g(x_i)^T g(x_i)}{g(x_i)^T H(x_i) g(x_i)}
      \]
    - This can serve as a good starting point for a backtracking line search.
Unconstrained Optimization

- Method of steepest decent:
  - Example: $f(x) = x_1^2 + 10x_2^2$
  - Contours for the function:

\[ x_{i+1} = x_i - \alpha_i g(x_i) \]

quadratic approximation:
Unconstrained Optimization

- Method of steepest decent:
  - Example: \( \log f(x) = x_1^2 + 10x_2^2 \)
  - Contours for the function:

\[
\alpha_i = \frac{g(x_i)^T g(x_i)}{g(x_i)^T H(x_i) g(x_i)}
\]

\[
x_{i+1} = x_i - \alpha_i g(x_i)
\]
10.34 Numerical Methods Applied to Chemical Engineering
Fall 2015

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.