12. Optimization

Outline

- Definition and examples
- Single variable (golden section search, Newton's method)
- Multi-variable (just Newton's)

Definition

- Given a known function, $f(x)$, find a point that is a local optimum (maximum or minimum, depending on problem)
- By convention, we usually consider minimization – note that $\max f(x) = -\min -f(x)$
- Functions may have many local optima – plotting the function can find a range over which the desired minimum is the only one

Applications

- Minimizing travel time
- Minimize cost in building, water distribution, ...
- How much to charge to maximize profit (homework problem)
- Solving regression problems (Section 10)

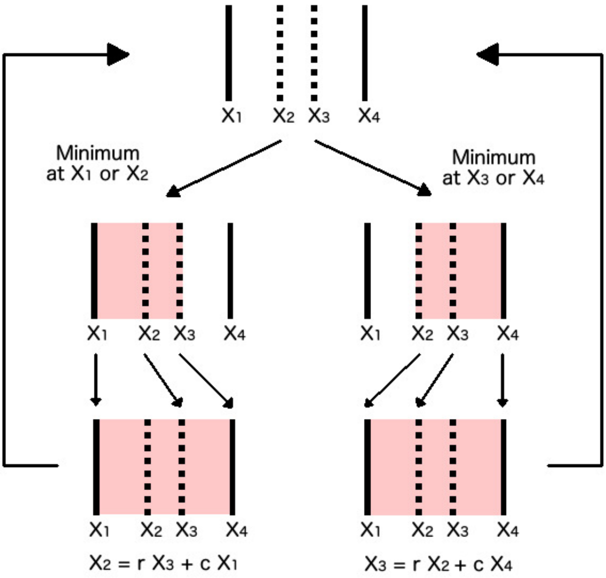
Examples from calculus

- What's the largest-volume open box that can be constructed from a 1.5 m by 1 m cardboard rectangle?
- What's the largest volume a cylinder with a surface area of 100 cm^2 can have?
- What size board with rectangular cross-section should be cut from a 2-m diameter cylindrical log to maximize its strength under cantilever loading, if this strength is proportional to xy^2, where x and y are the rectangle sides?
- You're designing a 3-m long hollow concrete column whose cost in dollars is given by $(m/m_0)^{1.2}$, where m is the column mass and m_0 is 1 kg. If your budget is \$500 and the density of concrete is 2000 kg m^{-3}, what diameter d and thickness t should the column have to be as strong as possible, given that the strength (against buckling) is proportional to $dt(d^2 + t^2)$?

Golden-section search

- Similar to bisection in that it will always find a minimum if one exists within an interval
- Converges at a slow, predictable rate
- Uses the 'golden ratio', $\phi = \frac{\sqrt{5}+1}{2} \approx 1.618$ ($\phi - 1 = \frac{1}{\phi}$)
- Each iteration:
  - Given an interval $[a, b]$
  - Compare the values of f at $c = a + (\phi - 1)(b - a)$ and $d = a + (2 - \phi)(b - a)$
  - If $f(c) < f(d)$, the new interval is $[b, d]$
  - If $f(d) < f(c)$, the new interval is $[a, c]$
  - Regardless, interval width drops by a factor of ϕ each iteration, so after n iterations, error decreases by about ϕ^{-n} (linear convergence)
- To find a maximum instead of a minimum, either replace f by $-f$ or replace $<$ by $>$

Illustration



Example

Find a minimum of $f(x) = x - 5\sin(x)$
By plotting the function, we see that it has a local minimum for x between 1 and 2
Choosing $a = 1, b = 2$, for the first iteration of golden section search we find that $c = 1.618, d = 1.382$ and $f(d) < f(c)$, so the new interval is 1 to 1.618
With $a = 1, b = 1.618$, we have $c = 1.382, d = 1.236$ and $f(c) < f(d)$, so the new interval is 1.618 to 1.236
With $a = 1.618, b = 1.236$, we have $c = 1.382, d = 1.472$ and $f(c) < f(d)$ still, so the new interval is 1.236 to 1.472
We would need about 30 iterations to find the minimum to 6 accurate digits ($\phi^{-29} \approx 10^{-6}$)

Newton's method

- Do root finding for the derivative of f, since at an optimum point, this derivative should be 0
- Works for finding a minimum or maximum
- Iteration is $x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$
- Converges fast if starting value is 'good'

Example

For this function, the iteration is $x_{i+1} = x_i - \frac{1-5\cos(x_i)}{5\sin(x_i)}$
If we start with $x_0 = 1$, we have
$x_1 = 1.4044, x_2 = 1.3695, x_3 = 1.3694$

Multi-dimensional optimization

- Often f is actually a function of several variables, which we can write as $f(\boldsymbol{x})$.
- Finding an optimum then is a more difficult problem
  - Golden-section search doesn't work in this case
  - Newton's method can be generalized to work
  - In practice, there are many numerical methods available – you'll learn more about them in CE 316

Newton's method for multidimensional optimization

- The equivalent of $f'(x)$ is the gradient vector of first partial derivatives, $\nabla(\boldsymbol{x})$, where $\nabla_i = \frac{\partial f}{\partial x_i}$
- The equivalent of $f''(x)$ is the (symmetric) Hessian matrix of second partial derivatives, $\boldsymbol{H}(\boldsymbol{x})$, where $H_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$
- The iteration is $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - \boldsymbol{H}^{-1}(\boldsymbol{x}_i)\nabla(\boldsymbol{x}_i)$, or (computationally preferable) solve $\boldsymbol{H}(\boldsymbol{x}_i)\Delta_i = \nabla(\boldsymbol{x}_i)$ for the vector Δ_i, and then $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - \Delta_i$

Example

If $f = x_1^2 + x_2^2 - x_1 x_2 - e^{-x_1^2} + 10\cos(x_1 - x_2)$, then

$$\nabla = \begin{pmatrix} 2x_1 - x_2 + 2x_1 e^{-x_1^2} - 10\sin(x_1 - x_2) \\ 2x_2 - x_1 + 10\sin(x_1 - x_2) \end{pmatrix}$$

$$\boldsymbol{H} = \begin{pmatrix} 2 + (2 - 4x_1^2)e^{-x_1^2} - 10\cos(x_1 - x_2) & -1 + 10\cos(x_1 - x_2) \\ -1 + 10\cos(x_1 - x_2) & 2 - 10\cos(x_1 - x_2) \end{pmatrix}$$

By plotting the function, we might come up with $\boldsymbol{x}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ as a good starting point for finding a minimum.

We then have $\nabla(\boldsymbol{x}_0) = \begin{pmatrix} -5.3572 \\ 6.0930 \end{pmatrix}$ and

$$\boldsymbol{H}(\boldsymbol{x}_0) = \begin{pmatrix} 5.4257 & -5.1615 \\ -5.1615 & 6.1615 \end{pmatrix}$$

Then $\boldsymbol{x}_1 = \boldsymbol{x}_0 - \boldsymbol{H}(\boldsymbol{x}_0) \setminus \nabla(\boldsymbol{x}_0) = \begin{pmatrix} 1.2297 \\ -1.7965 \end{pmatrix}$.

More steps

Repeating the process results in $\boldsymbol{x}_2 = \begin{pmatrix} 0.9533 \\ -1.7421 \end{pmatrix}$ and $\boldsymbol{x}_3 = \begin{pmatrix} 0.9608 \\ -1.7242 \end{pmatrix}$.

We can check that we are actually moving toward a minimum by monitoring f. $f(\boldsymbol{x}_0) = -1.5293, f(\boldsymbol{x}_1) = -3.2053, f(\boldsymbol{x}_2) = -3.8195, f(\boldsymbol{x}_3) = -3.8203$, which is the direction we want.