Unconstrained Optimization

CE 377K

January 27, 2015

ANNOUNCEMENTS

Exam date and ITE conference

REVIEW

Example problems: what were the “tricks”?

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & \sum_{a \in A} t_a x_a \\
\text{s.t.} \quad & \sum_{a \in A(i)} x_a - \sum_{a \in B(i)} x_a = \begin{cases} 1 & \text{if } i = r \\ -1 & \text{if } i = s \\ 0 & \text{otherwise} \end{cases} & \forall i \in \{1, \ldots, I\} \\
& x_a \in \{0, 1\} & \forall a \in \{1, \ldots, A\}
\end{aligned}
$$

Optimization problems from your own life?

SOME DEFINITIONS AND USEFUL FACTS

A feasible solution $\mathbf{x}^* \in X$ is a *global minimum* of f if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all feasible $\mathbf{x}$, and a *global maximum* if $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all feasible $\mathbf{x}$

An *optimal* solution is a global minimum (for a minimization problem) or a global maximum (for a maximization problem).

It is easy to convert back and forth between maximization and minimization problems. If the feasible set is X, the feasible solution $\mathbf{x}^*$ is a global maximum of f iff it is a global minimum of $-f$.

As a result, we do not need to develop different techniques for maximization or minimization problems. **The default convention in this class is to work with minimization problems.**

Other useful facts:

- Constants can be added or subtracted to an objective function without changing the optimal solutions.
- You can multiply an objective function by a nonnegative constant without changing the optimal solutions.

UNCONSTRAINED OPTIMIZATION: ONE DIMENSION

To start solving optimization problems, first consider optimization problems with no constraints at all.

In one dimension, the optimization problem is simply $\min_x f(x)$

This was the kind of problem you saw in calculus.

Assume that f is differentiable (and therefore continuous).

A *stationary point* of f is a value x such that $f'(x) = 0$.

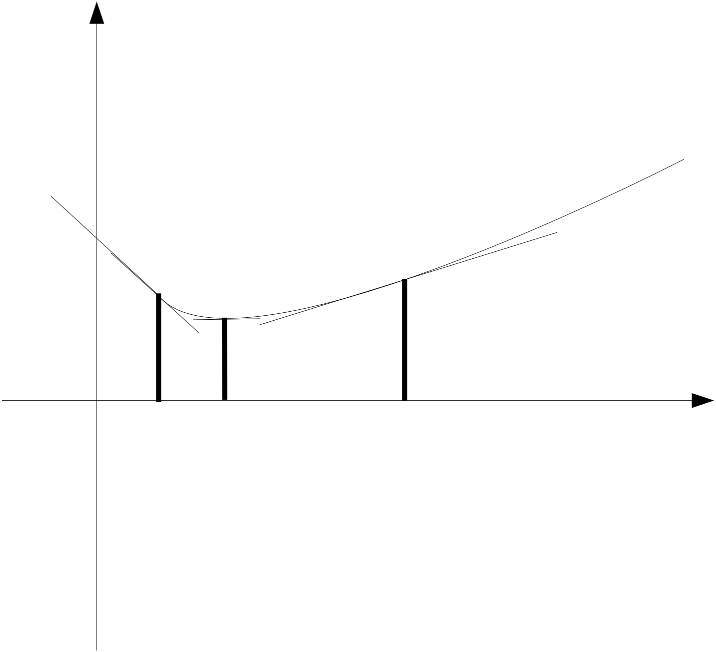
A *local minimum* of f is a value x^* such that $f(x^*) \leq f(x)$ for all x in some open interval (ℓ, h) containing x^*.

(Remember that a *global minimum* is a value x^* such that $f(x*) \leq f(x)$ for *all* x.)

Every global optimum of f is a local optimum, and if f is differentiable every local optimum is a stationary point.

Proof sketch. The first half is trivial.

For the second half, argue by contradiction. What if f was differentiable, but there was some local optimum which was *not* a stationary point?



This result does *not* hold in the opposite direction.

Some examples:

- $f(x) = -x^2$
- $f(x) = x^3$

So, in general, we can't hope to find the global minimum by enumerating all stationary points and seeing which has the least value.

However, with some additional assumptions on f this approach can work. One condition (there are others) is that f is *coercive*.

f is *coercive* if $f(x) \to +\infty$ as $|x| \to \infty$.

If f is differentiable and coercive, then f has a global minimum, which is the stationary point with least value of f.

Proof sketch. Differentiability and coercivity imply existence of a global minimum. Global minima must be local minima, which must be stationary points.

Examples:

- $f(x) = x^2 - 3x + 5$
- $f(x) = x^4 - 2x^2$
- $f(x) = e^x + e^{-x}$

UNCONSTRAINED OPTIMIZATION: MULTIPLE DIMENSIONS

Similar techniques apply when there is more than one decision variable and $\mathbf{x}$ is a vector.

The equivalent of the derivative is the *gradient* ∇f, the vector of all first partial derivatives:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) & \frac{\partial f}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

A stationary point is a point where the gradient is the zero vector (*all* partial derivatives equal zero)

A coercive function is one that tends to $+\infty$ as $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \to \infty$

The same results from the one dimensional case transfer to the multi-dimensional case:

- Every global minimum is a local minimum.
- If f is differentiable, then every local minimum is a stationary point.
- If f is coercive, then a global minimum exists.
- If f is differentiable and coercive, the global minimum is the stationary point with least objective value.

Examples:

- $f(x_1, x_2) = x_1^2 + x_2^2$
- $f(x_1, x_2) = 2x_1^2 + x_2^2 - x_1 x_2 - 7x_2$
- $f(x_1, x_2) = \frac{1}{2}x_1^2 - \frac{1}{2}x_2^2 + x_1 x_2 - 2x_1$
- $f(x_1, x_2) = -x_1 x_2 \exp\left(-\frac{1}{2}(x_1^2 + x_2^2)\right)$