Brief introduction to Lagrange multipliers

COMP90051, SM2 2015

Andrey Kan

In some machine learning methods one has to deal with problems such as find

$$x^* = \arg\min_x f(x), \text{ subject to } g(x) \geq 0$$

In Lagrange multiplier method, this problem is addressed with the use of auxiliary function called Lagrangian, and additional parameter λ called Lagrange multiplier:

$$L(x, \lambda) \stackrel{\text{def}}{=} f(x) - \lambda g(x)$$

One can then finds an infimum at any fixed λ, which yields a function called Lagrange dual:

$$L_D(\lambda) \stackrel{\text{def}}{=} \inf_x L(x, \lambda)$$

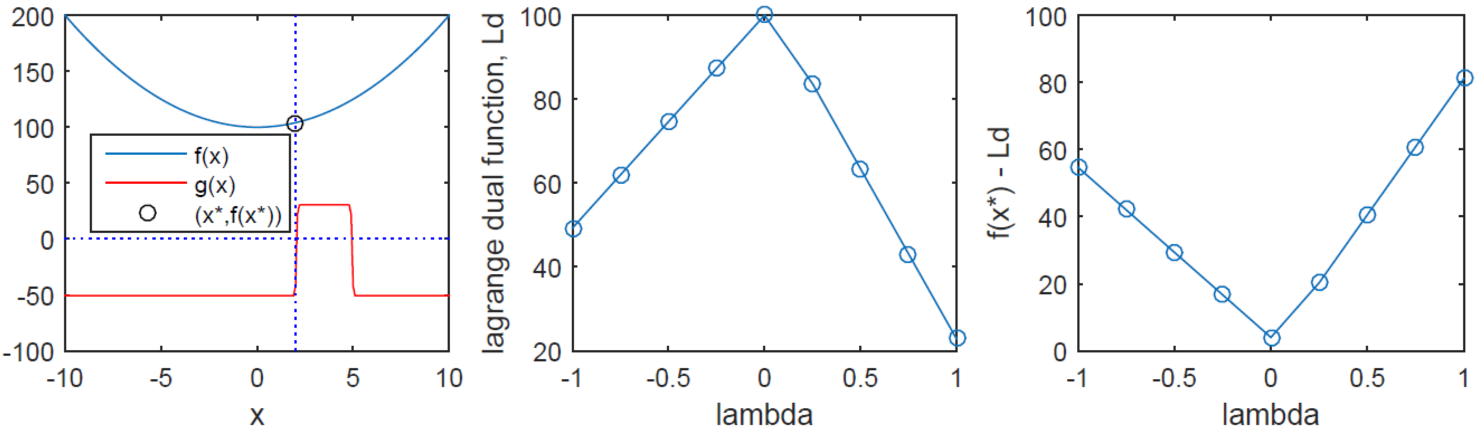
One can then consider the difference $f(x^*) - L_D(\lambda)$. This difference is always non-negative.

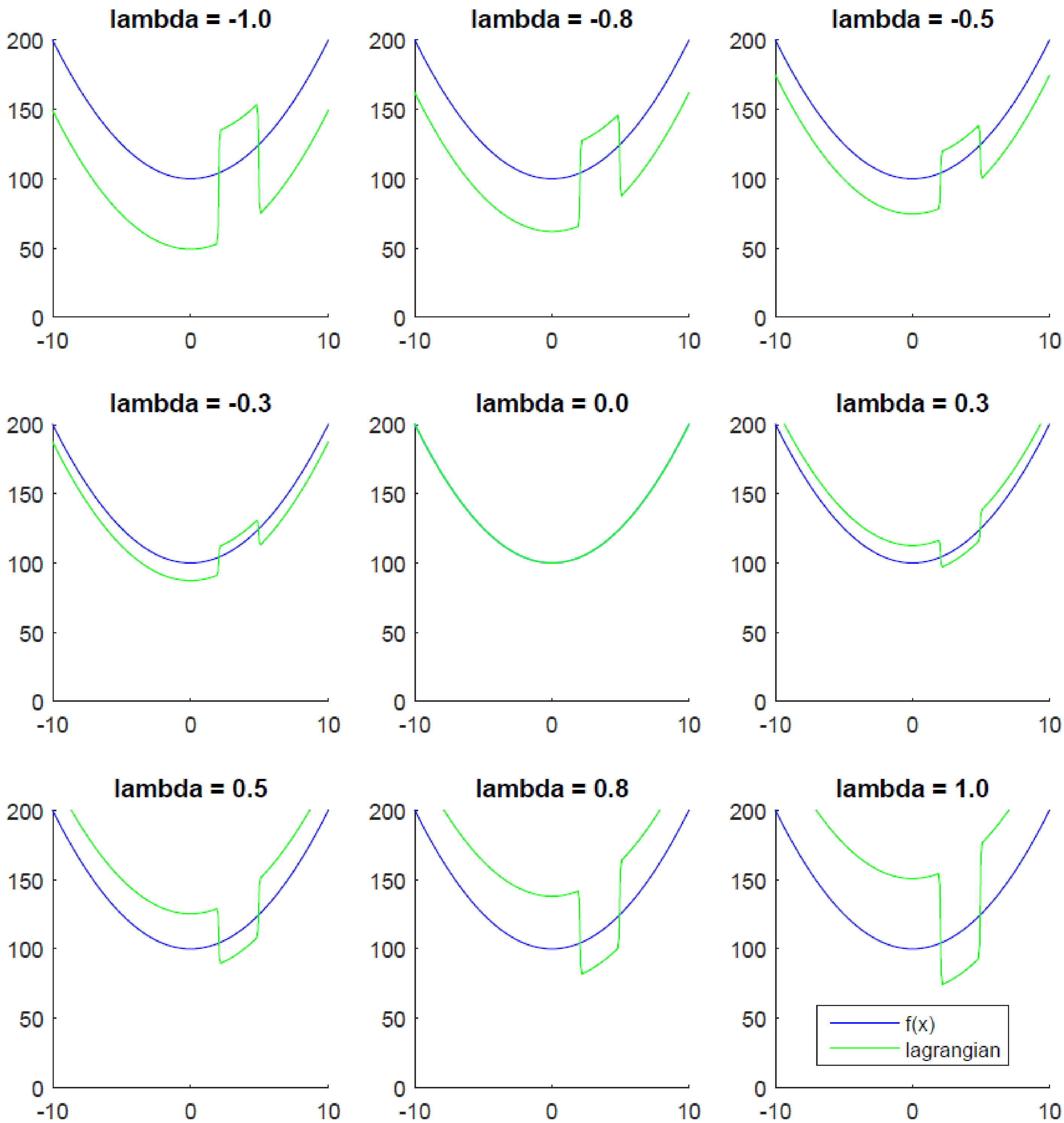
The aim is then to minimize this difference $\lambda^* = \arg\max_\lambda L_D(\lambda)$

Duality gap is then $f(x^*) - L_D(\lambda^*)$.

For some problems this gap is zero.

These concepts are illustrated below. (The gap is also zero in this example, but we are making too coarse of a step for λ.)




lambda = -1.0
lambda = -0.8
lambda = -0.5
lambda = -0.3
lambda = 0.0
lambda = 0.3
lambda = 0.5
lambda = 0.8
lambda = 1.0
200
150
100
50
0
-10
0
10
f(x)
lagrangian