Introduction to optimization and operations research
Interactive session

Michel Bierlaire

December 6, 2024

The correct answer is identified by the bold font.

1 Discrete optimization

1.1 Bounds

Consider a minimization problem P

$$\min_{x \in \mathbb{R}^n} f(x)$$

subject to

$$\begin{aligned} g(x) &\leq 0, \\ h(x) &= 0, \\ x &\geq 0, \\ x &\in X \subseteq \mathbb{R}^n. \end{aligned}$$

Consider x^* an optimal solution of this problem. For each of the following methods, decide if it provides a lower bound of $f(x^*)$, an upper bound, or none.

1. Consider the problem

   $$\min_{x \in \mathbb{R}^n} f(x)$$

   subject to

   $$\begin{aligned} g(x) &\leq 0, \\ h(x) &= 0, \\ x &\geq 0, \end{aligned}$$

and x_1 its optimal solution. Then, $f(x_1)$ is

(a) **a lower bound.**

(b) an upper bound.

(c) neither an upper or a lower bound.

It is a lower bound on $f(x^*)$. This is exploited by the branch and bound algorithm, where $X = \mathbb{Z}^n$. The relaxation of an integer linear optimization problem is solved with the simplex algorithm to obtain a lower bound.

2. Consider a feasible point x_2. Then, $f(x_2)$ is

   (a) a lower bound.

   (b) **an upper bound.**

   (c) neither an upper or a lower bound.

   It is an upper bound on $f(x^*)$, by definition of the optimal solution of a minimization problem.

3. Consider the problem

$$\min_{x \in \mathbb{R}^n} L(x; \lambda) = f(x) + \lambda^T h(x)$$

   subject to

$$\begin{aligned} g(x) &\leq 0, \\ x &\geq 0, \\ x &\in X \subseteq \mathbb{R}^n. \end{aligned}$$

   and x_3 its optimal solution. Then, for any λ, $L(x_3; \lambda)$ is

   (a) **a lower bound.**

   (b) an upper bound.

   (c) neither an upper or a lower bound.

   It is a lower bound on $f(x^*)$. L is the Lagrangian.

$$L(x_3; \lambda) \leq L(x^*; \lambda) = f(x^*) + \lambda^T h(x^*) = f(x^*).$$

   Note that x_3 may not be feasible for the original problem.

4. Consider the origin of the coordinate system, $x_4 = 0$. Then, $f(x_4)$ is

(a) a lower bound.

(b) an upper bound.

(c) **neither an upper or a lower bound.**

It is not a bound on $f(x^*)$ in the general case. If x_4 happens to be feasible, it provides an upper bound. If not, nothing can be said.

2 Non linear optimization

2.1 Newton local

Let f be a twice differentiable function. Which statement about Newton's local method for the minimization of f is **correct**?

Remember that the quadratic model is defined as

$$m_{x_k}(d) = f(x_k) + d^T \nabla f(x_k) + \frac{1}{2} d^T \nabla^2 f(x_k) \tag{1}$$

and Newton's equation is defined as

$$\nabla^2 f(x_k) d = -\nabla f(x_k). \tag{2}$$

1. **If the algorithm converges, it always converges to a stationary point of the function f.** Indeed, a stationary point x^* is such that $\nabla f(x^*) = 0$, which is exactly the set of equations that Newton's method is solving.

2. The point obtained during the k^{th} iteration of the algorithm maximizes the quadratic model of the function f in point x_k. No. It minimizes it, if the model is convex. If the function is concave, either it maximizes it (if Newton's equation (2) is solved) or it fails (if the quadratic model (1) is minimized).

3. If we start the algorithm from two different starting points, it will always converge to two different local minima. No. If x^* is a local mimimum, there is a neighborhood around x^* such that the method converges to x^* if started from any point within this neighborhood.

4. If the algorithm converges, it enables to always find a point that satisfies the second order necessary optimality condition. No, the method may converge to a stationary point that is a saddle point or a maximum, where the necessary optimality conditions are not verified.

2.2 Newton local

We want to minimize the function $f(x)$. We consider the iterate x_k of the local Newton method, such that the function f is not convex at x_k. The method builds a quadratic model of f at x_k, and minimizes this model to obtain x_{k+1}. Which statement is **correct**?

1. $f(x_{k+1}) > f(x_k)$.
2. **The iteration is unsuccessful.**
3. The iteration is successful, but we cannot say if $f(x_{k+1}) > f(x_k)$ or $f(x_{k+1}) < f(x_k)$.
4. $f(x_{k+1}) < f(x_k)$.

As the function is not convex at x_k, the quadratic model is not bounded from below. Therefore, there is no minimum, and the next iterate is not defined. The iteration is unsuccessful.

2.3 Line search

Consider a function f, a point x_k such that $\nabla f(x_k) \neq 0$, and a descent direction d_k at x_k. Let α^* be the step that minimizes the function f in the direction d_k:

$$\alpha^* = \operatorname{argmin}_\alpha f(x_k + \alpha d_k).$$

Which statement is **wrong**?

1. $f(x_k + \alpha^* d_k) < f(x_k)$. This is correct, from the definition of α^*, and the fact that d_k is a descent direction.

2. **The first Wolfe condition is verified at α^* for any β_1 between 0 and 1**:

   $$f(x_k + \alpha^* d_k) \leq f(x_k) + \alpha^* \beta_1 \nabla f(x_k)^T d_k \,,$$

   for all $0 < \beta_1 < 1$. This is incorrect. It depends on the value of β_1. If it is close to 1, the condition may be quite conservative, and the optimal step α^* may be deemed too long.

3. The second Wolfe condition is verified at α^* for any β_2 between 0 and 1

   $$\nabla f(x_k + \alpha^* d_k)^T d_k \geq \beta_2 \nabla f(x_k)^T d_k$$

   for all $0 < \beta_2 < 1$. Indeed, as α^* is the minimum of the function f along d_k, we have

   $$\nabla f(x_k + \alpha^* d_k)^T d_k = 0.$$

   Therefore, the second Wolfe condition at α^* is

   $$\beta_2 \nabla f(x_k)^T d_k \leq 0.$$

   As d_k is a descent direction and $\beta_2 > 0$, this is always verified.

2.4 Preconditioner

Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ and an iterate x_k such that the function is not convex at x_k. Among the following matrices, which one can be used to precondition the gradient?

1. $D_k = \nabla^2 f(x_k)$. No. As the function is not convex, $\nabla^2 f(x_k)$ is not positive definite.
2. $D_k = \frac{1}{2}\nabla^2 f(x_k)$. No, for the same reason as above.
3. $D_k = \nabla^2 f(x_k)^{-1}$. No. As $\nabla^2 f(x_k)$ is not positive definite, the same is true for its inverse.
4. $D_k =$
$$\begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}$$
Yes. It is a positive definite matrix.
5. $D_k =$
$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$$
No. It is a singular matrix.
6. $D_k =$
$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$
No. It is not a positive definite matrix.