Unconstrained Optimization

Not Open

1. Let A be **compact**, not void subset of $(\mathbb{R}^d, ||\cdot||)$ and f continious, then

$$\exists x^* \in A \text{ s.t. } f(x^*) = \min_{x \in A}[f(x)]$$

2. Let A be **closed**, non void subset of $(\mathbb{R}^d, ||\cdot||)$ and functions $f : A \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ s.t.

$$f(x) \geq g(||x||) \quad \forall x \in \mathbb{R}^d$$

$$\lim_{t \to +\infty} g(t) = +\infty$$

f is infinite at infinity if $\lim_{x \in A, ||x|| \to +\infty} f(x) = +\infty$

In dimension 1, to prove that f is infinite at infinity, it is sufficient to prove that

$$\lim_{|x| \to \infty} f(x) = +\infty$$

$$<=> \begin{cases} \lim_{x \to \infty} f(x) = +\infty \\ \lim_{x \to -\infty} f(x) = +\infty \end{cases}$$

Examples $\mathbb{R}^3, f(x,y,z) = \sqrt{x^2+y^2+z^2}, g(t) = t, => f(x,y,z) = ||x,y,z|| = g(||x,y,z||)$
$\lim_{t \to \infty} g(t) = +\infty => f$ is infinite at infinity.

$\mathbb{R}^2, f(x,y) = e^{-(x^2+y^2)} = e^{-(||x,y||^2)}, g(t) = e^{-t^2}, =>$
$f(x,y) = g(||x,y||)$
$\lim_{t \to \infty} g(t) = 0 => f$ is not infinite at infinity.

3. If A is **unbounded**, **closed**, non void subset and f is continious and **infinite at infinity**, then $\exists x^* \in A$, **global minimum** of A:

$$\min_{x \in A}[f(x)] = f(x^*)$$

Open

1. Find critical points: $\nabla f(x) = 0$
2. Check if Hessian is positive definite => **local** minimum
3. Check if function is convex $\iff$ Hessian is positive => **global** minimum
4. Check if function is strictly convex $\iff$ Hessian is positive definite => **unique** global minimum

5. If C is open in $\mathbb{R}^d$ and $f \in C^1(C, \mathbb{R})$, and f is **convex**, then

$$\forall x^* \in C, \nabla f(x^*) = 0 \iff f(x^*) = \min_{x \in C} f(x)$$

Constrained Optimization

1. A point is **admissible** if it belongs to the set of admissible points: $x \in A, A = \{x \in D : g_i(x) \leq 0, i = 1, \ldots, m$ and $h_j(x) = 0, j = 1, \ldots, p\}$. Look if at x^* all $g_i(x^*) \leq 0$ and $h_j(x^*) = 0$
2. Constraints are **qualified** at a point if:
   1. $\{\nabla h_j(x), j = 1, \ldots, p\} \cup \{\nabla g_i(x), i \in I(x)(g_i(x) = 0)\}$ are **linearly independent**.
   2. If tangent cone consists of directions: $T_x(A) = \{d \in R^d : \forall j \in \{1, ..., p\}, < \nabla h_j(x), d >= 0 \, \forall i \in I(x), < \nabla g_i(x), d > \leq 0\}$.
   3. Use **Slater** condition (if convex):
      1. Check if problem is convex:
         1. $f(x)$ is convex
         2. $g_i(x), i = 1, \ldots, m$ are convex
         3. $h_j(x) = a_j^T x - b_j, j = 1, \ldots, p, (h_1(x), \ldots, h_p(x)) = 0 \iff Ax = b$ are affine.
      2. Check if Slater condition holds:
         1. if $\exists x \in D$ s.t. $g_i(x) < 0, i = 1, \ldots, m$ and $h_j(x) = 0, j = 1, \ldots, p(Ax = b)$, then constraints are qualified at x.
         2. if $\exists k \in \{1, \ldots, m\}$ s.t. $g_1, \ldots, g_k$ **are affine**, then $\exists x \in D$ s.t. $g_i(x) \leq 0, i = 1, \ldots, k$, $g_i(x) < 0, i = k+1, \ldots, m$ and $h_j(x) = 0, j = 1, \ldots, p(Ax = b)$, then constraints are qualified at x
      3. If g_i's are affine and f is convex, then $\exists$ unique solution to (P).
3. Constraints are **qualified** at **all points**:
   1. Check if **constraints** are **linearly independent**.
4. **Tangent cone** to A at x^*: $T_{x^*}(A) = \{d \in R^d : \forall j \in \{1, ..., p\}, < \nabla h_j(x^*), d >= 0, \forall i \in I(x^*), < \nabla g_i(x^*), d > \leq 0\}$.
5. **Lagrangian**: $\mathcal{L}(x, \gamma, \lambda) = f(x) + \sum_{i=1}^m \gamma_i g_i(x) + \sum_{j=1}^p \lambda_j h_j(x)$
6. **KKT** critical points: if x^* is a local minimum of f on A, then $\exists(\gamma^*, \lambda^*) \in R^m \times R^p$ s.t.:
   1. $x^* \in A$ admissible, constrains are qualified at x^*

2. $\nabla_x \mathcal{L}(x^*, \gamma^*, \lambda^*) = 0$
3. $h_j(x^*) = 0$
4. $\gamma_i^* g_i^*(x^*) = 0$
5. $\gamma_i^* \geq 0$
6. $j = 1, ..., p$
7. $i = 1, ..., m$

7. If (P) is convex, x^* is admissible, constraints at x^* are qualified and KKT conditions hold at x^*, then x^* is **global minimum**.

8. **Hessian of the Lagrangian**: $H\mathcal{L}(x, \gamma, \lambda) = Hf(x) + \sum_{i=1}^{m} \gamma_i Hg_i(x) + \sum_{j=1}^{p} \lambda_j Hh_j(x)$

9. If x^* is admissible, constraints at x^* are qualified, KKT conditions hold at x^* and HL is positive definite on $T_{x^*}^+(A)$ (where x^* is qualified by linear independence), then x^* is a **strict local minimum**.

10. **Lagrange dual function**: $\phi(\gamma, \lambda) = \inf_{x \in D} L(x, \gamma, \lambda)$, $\inf_{x \in D} L(x, \gamma, \lambda) = -\sup_{x \in D}[-L(x, \gamma, \lambda)]$

11. **Dual & primal problems**: (P) : $\min_{x \in A} f(x) = p^* -$ primal problem, (D) : $\max_{\gamma \in R^m, \lambda \in R^p} \phi(\gamma, \lambda) = d^* -$ dual problem, $d^* = \sup_{\gamma \in R^m, \lambda \in R^p} \phi(\gamma, \lambda) \leq \inf_{x \in A} f(x) = p^*$, $d^* \leq p^*$

12. **Conjugate function**: $f^*(y) = \sup_{x \in D(f)}(y^T x - f(x))$

13. **Saddle point**: $(x^*, \gamma^*, \lambda^*)$:
    1. $L(x^*, \gamma, \lambda) \leq L(x^*, \gamma^*, \lambda^*), \forall \gamma \in R_+^m, \lambda \in R^p$ (maximizes)
    2. $L(x^*, \gamma^*, \lambda^*) \leq L(x, \gamma^*, \lambda^*), \forall x \in D$ (minimizes)

14. **Zero duality gap**: if x^* is solution of (P), (γ^*, λ^*) solution of (D) and $(x^*, \gamma^*, \lambda^*)$ is saddle point, then $p^* = d^*$

15. If x^* is admissible, constraints are satisfied at x^*, KKT conditions hold at x^* with some (γ^*, λ^*), then x^* is **global minimum** if $(x^*, \gamma^*, \lambda^*)$ is a saddle point.

16. If (P) has an optimal solution, so does (D) and $p^* = d^*$.

17. We have: $f(x^*) = L(x^*, \gamma^*, \lambda^*) = \inf_{x \in R^d} L(x, \gamma^*, \lambda^*) \iff x^* \in A, \gamma \geq 0$ and $x^*, (\gamma^*, \lambda^*)$ are solutions of $(P), (D)$, respectively.

18. If x^* is solution of (P), (P) is convex and Slater's conditions are satisfied, then $p^* = d^*$ and $\exists (\gamma^*, \lambda^*) \in R^m \times R^p$ s.t. $(x^*, \gamma^*, \lambda^*)$ is a saddle point of the Lagrangian and $\nabla f(x^*) + \sum_{i=1}^{m} \gamma_i^* \nabla g_i(x^*) + \sum_{j=1}^{p} \lambda_j^* \nabla h_i(x^*) = 0$

Algorithm to Find Local Minimum

1. Find **admissible** points.
2. Check if constraints are **qualified** at admissible points.
3. Check if (P) is **convex**.
    1. If convex => x^* is **global minimum**.
    2. If not convex:
        1. Check if KKT conditions hold.
            1. If HL is positive definite on $T_{x^*}^+(A)$ (where x^* is qualified by linear independence), then x^* is a **strict local minimum**
            2. If HL is not semi-positive definite, x^* cannot be a local minimum.
            3. If $(x^*, \gamma^*, \lambda^*)$ is saddle, then x^* is a **global minimum**.