BUSINESS OPTIMIZATION
AND SIMULATION

Module 4
Nonlinear optimization
STRUCTURE OF THE MODULE

• Sessions:
  • The unconstrained case: formulation, examples
  • Optimality conditions y methods
  • Constrained nonlinear optimization problems
  • Solving constrained problems
UNCONSTRAINED PROBLEMS

• Consider the case of optimization problems with no constraints on the variables and a nonlinear objective function

\[ \min_x f(x) \]

• Local solutions: cannot be improved by values close to the solution
• Global solutions: the best of all local solutions
OPTIMIZING UNCONSTRAINED PROBLEMS

• Ideal outcome: compute a global solution
  • In general, it is not possible to find one in a reasonable amount of time
• Two alternatives:
  • Accept a quick local solution
  • Attempt to compute a heuristic approximation to a global solution
    • Or one based on deterministic algorithms if the dimension of the problem is not large
• In general, basic versions of optimization solvers are only able to find local solutions
  • Global solutions only found when some conditions are satisfied:
    • Convexity: local optimizers = global optimizers
OPTIMIZING UNCONSTRAINED PROBLEMS

- Advanced solvers use heuristics to compute approximations to the global optimizers under general conditions
  - Without imposing requirements on differentiability, convexity, etc.
- In practice:
  - If we maximize and the function is concave, we may obtain the maximizer in a reasonable amount of time
  - If we minimize and the function is convex, we may compute the minimizer in a reasonable amount of time
EXAMPLE 1: SMARTPHONE MARKETING

• Description:

  • A company wishes to sell a smartphone to compete with other high-end products
    • It has invested one million euros to develop this product
    • The success of the product will depend on the investment on the marketing campaign and the final price of the phone

  • Two important decisions:
    • $a$ : amount to invest in the marketing campaign
    • $p$ : price of the smartphone
EXAMPLE 1: SMARTPHONE MARKETING

• Description:

• Formula used by the marketing department to estimate the sales of the new product during the coming year:

\[ S = 20000 + 5\sqrt{a} - 60p \]

• The production cost of the phone is 100 euros/unit

• How could the company maximize its profits for the coming year?
EXAMPLE 1: SMARTPHONE MARKETING

• Model:
  • Profits from sales: 
    \[(20000 + 5\sqrt{a} - 60p)p\]
  • Total production costs: 
    \[(20000 + 5\sqrt{a} - 60p)100\]
  • Development costs: 
    \[1000000\]
  • Marketing costs: 
    \[a\]
  • Total profit: 
    \[(20000 + 5\sqrt{a} - 60p)(p - 100) - 1000000 - a\]
EXAMPLE 1: SMARTPHONE MARKETING

- Optimal strategy?
  - Maximize profit
- Constraints?
  - Nonnegative values for the variables
    - Do you need to include them?
- Initial iterate:
  - What happens if the initial values are negative?
  - What if they are large and positive?
  - Small and positive?
- Is the problem convex?
  - Does the problem have more than one local solution?
  - Can you compute the global solution?
EXAMPLE 2: DATA FITTING

• Regression problems

• How to fit a model to some available data

• Different approaches: criteria to define what is best

• Least squares:

\[
\min_{\beta} \frac{1}{2} \sum_i (y_i - x_i^T \beta)^2
\]

• Nonlinear least squares:

\[
\min_{\beta} \frac{1}{2} \sum_i (y_i - F_i(\beta; x_i))^2
\]

• Minimum absolute deviation:

\[
\min_{\beta} \sum_i |y_i - x_i^T \beta|
\]
EXAMPLE 2: DATA FITTING

- An specific example: exponential or logit regression
  
- For example, it may be of interest to study the relationship between the growth rate of a person and his/her age

  - This relationship is nonlinear

  - The rate is high in the first years of life and then it stabilizes

- A model could be

\[
\text{rate} = \beta_0 + \beta_1 \exp(\beta_2 \text{age}) + \text{error}
\]
UNCONSTRAINED OPTIMALITY CONDITIONS

- When solving practical problems:
  - We may fail to obtain a solution
    - We need good estimates for the initial values of the variables
  - Even if we find a solution, in many cases we have no information about other possible solutions
    - Try with different starting points
- How can we obtain better information about the solutions?
  - Theoretical properties
    - Study the conditions satisfied at a solution
      - Check if they are satisfied
    - Or use them to find other candidate solutions
UNCONSTRAINED OPTIMALITY CONDITIONS

- Unconstrained optimization problem: $$\min_{x} f(x)$$

- If we wish to maximize the objective function, we could also solve $$\min_{x} -f(x)$$

- A point (or a decision) $$x^*$$ is a local solution if there is no better alternative close to it
  $$\exists \epsilon > 0, f(x^*) \leq f(x) \quad \forall x : \|x - x^*\| < \epsilon$$

- A point (or a decision) $$x^*$$ is a global solution if there is no better point (in all the space)
  $$f(x^*) \leq f(x) \quad \forall x$$
OPTIMALITY CONDITIONS

- Necessary conditions:

- Univariate case:
  \[ f'(x) = 0, \quad f''(x) \geq 0 \]

- Extension to the multivariate case

- First-order conditions:
  - If \( x^* \) is a local minimizer, then

- Second-order conditions:
  - If \( x^* \) is a local minimizer, then
OPTIMALITY CONDITIONS

• Sufficient conditions:

  • Univariate case:

  \[ f'(x) = 0, \quad f''(x) > 0 \]

  • Extension to the multivariate case

  • If the following conditions hold at \( x^* \), it is a local minimizer:

  \[ \nabla f(x^*) = 0 \]

  \[ \nabla^2 f(x^*) \geq 0 \]
OPTIMALITY CONDITIONS

• Example:
  
  • Consider the unconstrained problem:

  \[
  \min_x f(x), \quad f(x_1, x_2) = \frac{1}{3}x_1^3 + \frac{1}{2}x_1^2 + 2x_1x_2 + \frac{1}{2}x_2^2 - x_2 + 9
  \]

  • Necessary conditions:

  \[
  \nabla f(x) = \begin{pmatrix} \frac{2}{3}x_1^2 + x_1 + 2x_2 \\ 2x_1 + x_2 \end{pmatrix} = 0
  \]

  • There exist two stationary points (minimizer candidates):

  \[
  x_a = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad x_b = \begin{pmatrix} 2 \\ -3 \end{pmatrix}
  \]
OPTIMALITY CONDITIONS

- Example:
  - Sufficient condition

  \[ \nabla^2 f(x) = \begin{pmatrix} 2x_1 + 1 & 2 \\ 2 & 1 \end{pmatrix} \Rightarrow \]

  \[ \nabla^2 f(x_a) = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \] and \[ \nabla^2 f(x_b) = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \]

  - Thus,

    - \( \nabla^2 f(x_b) \) is positive definite \( \Rightarrow x_b \) local minimizer
    - \( \nabla^2 f(x_a) \) is indefinite, and \( x_a \) is neither a local minimizer nor a local maximizer
EXAMPLE 1: SMARTPHONE MARKETING

- A company wishes to sell a smartphone to compete with other high-end products

- Optimization model:

\[ (20000 + 5\sqrt{a} - 60p)(p - 100) - 1000000 - a \]

- First-order conditions:

\[
\nabla f = \begin{pmatrix}
\frac{5(p - 100)}{2\sqrt{a}} - 1 \\
20000 + 5\sqrt{a} - 120p + 6000
\end{pmatrix} = 0
\]

- One solution: \( a = 106003.245 \), \( p = 230.233 \)

- Second-order condition: Hessian matrix

\[
\nabla^2 f = \begin{pmatrix}
-\frac{5}{4}(p - 100)a^{-3/2} & \frac{5}{2}a^{-1/2} \\
\frac{5}{2}a^{-1/2} & -120
\end{pmatrix}
\]
OPTIMALITY CONDITIONS

• What happens if a minimizer does not satisfy the sufficient conditions?

\[ f_1(x) = x^3, \quad f_2(x) = x^4, \quad f_3(x) = -x^4 \]

• For all these functions it holds that \( \nabla f(0) = \nabla^2 f(0) = 0 \)

• Thus, \( x = 0 \) is a candidate for a local minimizer in all cases
  
  • But while \( f_2 \) has a local minimum at \( x = 0 \)
  
  • \( f_1 \) has a saddle point at that point
  
  • \( f_3 \) has a local maximum at the point

• The points satisfying these conditions are known as singular points
OPTIMALITY CONDITIONS

• Summary:

• A point is stationary if $\nabla f(x^*) = 0$

• For these points:
  
  • $\nabla^2 f(x^*) > 0 \Rightarrow$ minimizer
  
  • $\nabla^2 f(x^*) < 0 \Rightarrow$ maximizer
  
  • $\nabla^2 f(x^*)$ indefinite $\Rightarrow$ saddle point
  
  • $\nabla^2 f(x^*)$ singular $\Rightarrow$ any of the above
NEWTON’S METHOD

• Computing a (local) solution:

  • Most algorithms are iterative and descending
    • They compute points with decreasing values of the objective function

  \[ x_0, x_1, x_2, \ldots \text{ such that } f(x_{k+1}) < f(x_k), \ k = 0, 1, 2, \ldots \]

• The main step is to compute a search direction, \( p_k \), to take us from \( x_k \) to \( x_{k+1} \)

• Newton's method. The iterations take the form:

  \[ x_{k+1} = x_k + p_k, \quad p_k = -\left(\nabla^2 f(x_k)\right)^{-1}\nabla f(x_k) \]
CONSTRAINED PROBLEMS

• If we allow constraints on the variables, the problem is now

\[
\min_x f(x) \\
\text{s.t. } c(x) \geq 0
\]

• We consider the case when either the objective function or the constraints are nonlinear functions

• The optimizers may have significantly different properties than those corresponding to unconstrained problems

• Local solution: belongs to the feasible region and it cannot be improved in a feasible neighborhood of the solution

• Global solution: belongs to the feasible region, and is the best of all local solutions
CONSTRAINED PROBLEMS

• Solution properties:
  • Differences with unconstrained problems
    • Identifying the active constraints at the solution can be as important as finding points with good gradient values
  • Differences with linear problems
    • Solutions do not need to be at vertices
  • Finding a local solution. Either
    • Transform the problem to one without inequality constraints, or
    • Find the correct active constraints at a solution
    • Efficient trial and error procedures
CONSTRAINED PROBLEMS

- Practical difficulties:
  - Local solutions
    - If the problem is not convex, the solution found by the Solver may only be a local solution (not a global one)
    - Very difficult to check formally
    - Heuristic: You can try to solve the problem from different starting points
  - Ill-defined functions
    - In some cases, the objective or constraint functions may not be defined in all points (square roots, power functions, logarithms)
    - Even if you add constraints to avoid these points, the algorithm may generate infeasible points in that region
    - Heuristic: start close enough to a solution
EXAMPLE I: PORTFOLIO OPTIMIZATION

• The problem:
  
  • You have $n$ assets in which you can invest a certain amount of money
    
    • To simplify the formulation, we will assume this amount to be 1
  
  • The random variable $R_i$ represents the return rate associated to each asset
  
  • Your goal is to find the proportions $x_i$ to invest in each of the assets
    
    • To maximize your return (after one period)
    
    • And to minimize your investment risk
EXAMPLE 1: PORTFOLIO OPTIMIZATION

• The model:
  • We wish to solve:

\[
\begin{align*}
\text{maximize}_x & \quad \sum_i R_i x_i \\
\text{subject to} & \quad \sum x_i = 1
\end{align*}
\]

• Is this problem well-defined?

• A well-defined version:

\[
\begin{align*}
\text{maximize}_x & \quad \sum_i r_i x_i, \quad r_i \equiv \mathbb{E}[R_i] \\
\text{subject to} & \quad \sum x_i = 1, \quad x_i \geq 0
\end{align*}
\]

• But, is this reasonable?
EXAMPLE 1: PORTFOLIO OPTIMIZATION

• A reasonable version (Markowitz model):

\[
\begin{align*}
\text{maximize}_x & \quad r^T x - \frac{1}{2} \gamma x^T S x \\
\text{subject to} & \quad \sum_i x_i = 1
\end{align*}
\]

where $S = \text{Var}(R)$ and $\gamma$ is a risk-aversion coefficient

• This model allows the construction of an efficient frontier (policies that, for a given return, have minimum variance)

• It is a quadratic problem
EXAMPLE 1: PORTFOLIO OPTIMIZATION

• Another reasonable alternative:

$$\begin{align*}
\text{minimize}_x & \quad \text{VaR}_\beta \left( -(R_1 x_1 + \cdots + R_n x_n) \right) \\
\text{subject to} & \quad \sum x_i = 1
\end{align*}$$

where $\text{VaR}_\beta$ is the Value-at-Risk (percentile) corresponding to a given $0 \leq \beta \leq 1$

• This is a nonlinear, nonconvex problem
  • How can you compute a solution?
  • Advanced techniques of nonlinear optimization
SOLVING CONSTRAINED PROBLEMS

- Studying local solutions for a constrained problem:
  
  \[
  \begin{align*}
  \text{minimize} & \quad f(x) \\
  \text{subject to} & \quad c_E(x) = 0 \\
  & \quad c_I(x) \geq 0
  \end{align*}
  \]

- Use the optimality conditions to obtain additional information

- Form of the optimality conditions in the constrained case

\[
\begin{align*}
\nabla_x f(x^*) - \nabla_x c(x^*) \lambda^* &= 0 \quad \text{stationarity} \\
c_I(x^*) \geq 0 \text{ and } c_E(x^*) = 0 \\
c_I(x^*)^T \lambda_I^* &= 0 \quad \text{complementarity} \\
\lambda_I^* &\geq 0 \quad \text{multiplier sign}
\end{align*}
\]

- We say that \( x^* \) is a stationary point if they hold for some \( \lambda^* \)
SOLVING CONSTRAINED PROBLEMS

• The preceding conditions are necessary but not sufficient

• First-order optimality conditions (no second derivatives)

• The vector \( \lambda \) is known as the vector of Lagrange multipliers

• Part of the Karush-Kuhn-Tucker (KKT) conditions

• Second-order condition:

\[
L(x, \lambda) = f(x) - \sum_j \lambda_j c_j(x)
\]

\[
Z \text{ matrix with columns forming a basis for } \{d : \nabla \hat{c}(x)d = 0\}
\]

where \( \hat{c} \) denotes the active constraints, \( \hat{c}(x) = 0 \)

\[
Z^T \nabla^2_{xx} L(x, \lambda) Z \succeq 0
\]
OPTIMALITY CONDITIONS

• Example:

minimize \( f(x) = (x_1 - 3/2)^2 + (x_2 - 5/4)^2 \)

subject to \( c(x) = \begin{pmatrix} 1 - x_1 - x_2 \\ 1 - x_1 + x_2 \\ 1 + x_1 - x_2 \\ 1 + x_1 + x_2 \end{pmatrix} \geq 0 \)

• Check that the point \((1,0)\) satisfies the necessary conditions
OPTIMALITY CONDITIONS

- The multiplier vector $\lambda^* = (3/4, 1/4, 0, 0)$ satisfies

\[
\begin{pmatrix}
-1 \\
-0.5 \\
\end{pmatrix} - \begin{pmatrix}
-1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 \\
\end{pmatrix} \begin{pmatrix}
\lambda_1^* \\
\lambda_2^* \\
\lambda_3^* \\
\lambda_4^* \\
\end{pmatrix} = 0
\]

\[
\begin{pmatrix}
1 & -1 & -0 \\
1 & -1 & +0 \\
1 & 1 & -0 \\
1 & 1 & +0 \\
0 & 0 & 0 \\
2 & 0 & 2 \\
2 & 0 & 2 \\
\end{pmatrix} \begin{pmatrix}
\lambda_1^* \\
\lambda_2^* \\
\lambda_3^* \\
\lambda_4^* \\
\end{pmatrix} \geq 0
\]

$\lambda^* \geq 0$