Short review on optimisation theory

Unconstrained problem

$$\min f(x)$$

- f: $R^N \rightarrow R$
- f: objective function
- x: control variable vector
- The optimal point can be found solving the following system of equations (stationary condition):

$$\nabla f(x) = [0] \qquad \begin{cases} \dfrac{\partial f}{\partial x_1} = 0 \\ \dfrac{\partial f}{\partial x_2} = 0 \\ \vdots \\ \dfrac{\partial f}{\partial x_N} = 0 \end{cases}$$

Comments

- The solution of the system of equations could be a minimum, a maximum or a saddle point
  - It is necessary to check the Hessian of the objective function to know which type of solution we found
- It is a very sample problem from the theory point of view, but it could present several computational problem to find the solution
  - Generally, the system of equation is made by non linear equations
  - Numerical methods are necessary to find the solution (for example, Gauss based methods or Newton based methods)
- The objective function have to be $f \in C_2(R^N)$
  - This approach is not applicable in case of discrete of binary variables

Constrained optimization

$$\begin{aligned}
&\min f(x)\\
&s.t.\\
&g(x) = [0]\\
&h(x) \le [0]
\end{aligned}$$

- To find the solution of this optimisation problem is necessary to built the Lagrangian function

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{M} \lambda_i \cdot g_i(x) + \sum_{j=1}^{L} \mu_j \cdot h_j(x)$$

- After that, it is necessary to write the, so called, KKT condition. The solution of the equation system given by the KKT condition provide the optimal point only if the problem is convex (necessary and sufficient conditions)
  - An optimisation problem is convex if the OF is convex and the feasibility region is convex
- In case of non-convex problem, the KKT conditions are only necessary but not sufficient
  - a sub-optimal point could be found

$$
\begin{cases}
\dfrac{\partial L(x,\lambda,\mu)}{\partial x_1} = \dfrac{\partial f(x)}{\partial x_1} + \sum_{i=1}^{M} \lambda_i \cdot \dfrac{\partial g_i(x)}{\partial x_1} + \sum_{j=1}^{L} \mu_j \cdot \dfrac{\partial h_j(x)}{\partial x_1} = 0 \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial x_2} = \dfrac{\partial f(x)}{\partial x_2} + \sum_{i=1}^{M} \lambda_i \cdot \dfrac{\partial g_i(x)}{\partial x_2} + \sum_{j=1}^{L} \mu_j \cdot \dfrac{\partial h_j(x)}{\partial x_2} = 0 \\
\vdots \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial x_N} = \dfrac{\partial f(x)}{\partial x_N} + \sum_{i=1}^{M} \lambda_i \cdot \dfrac{\partial g_i(x)}{\partial x_N} + \sum_{j=N}^{L} \mu_j \cdot \dfrac{\partial h_j(x)}{\partial x_N} = 0 \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \lambda_1} = g_1(x) = 0 \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \lambda_2} = g_2(x) = 0 \\
\vdots \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \lambda_M} = g_M(x) = 0 \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \mu_1} = h_1(x) \le 0 \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \mu_2} = h_2(x) \le 0 \\
\vdots \\
\dfrac{\partial L(x,\lambda,\mu)}{\partial \mu_L} = h_L(x) \le 0 \\
\left.\begin{aligned}
&\mu_1 \cdot h_1(x) = 0 \\
&\mu_1 \ge 0 \\
&\mu_2 \cdot h_2(x) = 0 \\
&\mu_2 \ge 0 \\
&\vdots \\
&\mu_L \cdot h_L(x) = 0 \\
&\mu_L \ge 0
\end{aligned}\right\} \text{Exclusion conditions}
\end{cases}
$$

KKT conditions

Comments

- The equation system compared to the case without constrains is constituted by **N+M+L** equations and variables, instead of **N**
- The variables are the our control variables of the original problem (**N**) plus the Lagrange multipliers λ (**M**) and μ **(**L**)** associated to respectively the equality and inequality constraints

The meaning of the Lagrange multipliers

- Only in the optimal point $\left(x^*, \lambda^*, \mu^*\right)$ is valid the following equation

$$\frac{\partial f\left(x^*\right)}{\partial g_i\left(x^*\right)} = -\lambda_i^*$$

- The derivative of the objective function f(x) with respect to the value of the equality constraint equals, unless the sign, the value that assumes the Lagrange multiplier
- This multiplier tells us how much the objective function would change for a limited change of the value of the equality constraint

Example

- We suppose that we have an equality constraint

$$3x^2 + 2x - 3 = 0 \quad \leftrightarrow \quad 3x^2 + 2x = 3$$

- and the Lagrange multiplier is equal to 0.002
- This means that if we change the constraint of a very little quantity ($3x^2 + 3x = 3.001$) the objective function change of a quantity equal to:

$$\Delta f = \lambda \Delta g = 0.002 \cdot 0.001 = 2 \cdot 10^{-6}$$

The exclusion conditions

- If in the optimal point an inequality constraint is not active (i.e., $h_j(x)<0$) the exclusion equation may be valid only if the associated Lagrange multiplier is null.
- Conversely, if the constraint is active ($h_j(x)=0$), i.e. the inequality constraint assumes the limit value, then the Lagrange multiplier will be different from zero.
  - In this case, in the solution point, the inequality constraint is equivalent an equality constraint
- It is necessary to highlight that remains unaltered the meaning of the Lagrange multiplier, since, if the inequality constraint is satisfied with a strict inequality ($h_j(x)<0$), a variation of its limit value does not prevent further improvements of the objective function: therefore the Lagrange multiplier is equal to zero, i.e., also a relaxation of this constraint does not lead to a further improvement of the objective function and then the derivative of f(x) with respect to the limit value of the bond must be zero

Example: λ-Dispatching problem

- We have a set of different generating units with different operating costs due to fuel consumption: we know the fuel consumption curve for each unit
- We have a contract to provide energy to a consumer for a given hour
- We want to provide this energy maximizing our profit. Fixed the selling price, maximize the profit means minimize the generation cost
- The maximum and minimum power constraints for each unit have to be take into account
- Also, we neglect the constraints introduced by the transmission network (transmission constraints)
- We assume that the transmission losses are negligible

Consumption curves of thermoelectric generators

- Consumption curve to produce an amount of energy per hour while maintaining a constant power in the hour. This is a convex curve

$$C(P) = C_0 + C_1 \cdot P + C_2 \cdot P^2 \quad [€/h]$$

C_0 [€/h] $\quad$ C_1 [€/MWh] $\quad$ C_2 [€/MW²h]

$$\underset{P_i}{Min} \quad \sum_{i=1}^{N_g} F_i(P_i)$$

Subject to (s.t.)

$$\sum_{i=1}^{N_g} P_i = P_T$$

Equability constraint: balance equation

Inequability constraint: Technical conditions

$$\underline{P}_i \leq P_i \leq \overline{P}_i \qquad i = 1, \ldots, \mathrm{N_g}.$$

Problem characteristics

- The problem is convex: the KKT conditions are sufficient condition
- The control variables are the power output of each generating unit: P_1 , P_2 , …, P_{Ng}
- The Lagrange function of the problem is:

$$L(P,\lambda,\mu)=\sum_{i=1}^{Ng} f_i(P_i)-\lambda\left(\sum_{i=1}^{Ng} P_i - C_g\right)-\sum_{i=1}^{Ng}\underline{\mu_i}\left(P_i-\underline{P_i}\right)-\sum_{i=1}^{Ng}\overline{\mu_i}\left(\overline{P_i}-P_i\right)$$

$$\frac{\partial L}{\partial P}=0 \quad \frac{\partial L}{\partial \lambda}=0 \quad \frac{\partial L}{\partial \mu}\leq 0 \quad \underline{\mu_i}\left(P_i-\underline{P_i}\right)=0 \quad \overline{\mu_i}\left(\overline{P_i}-P_i\right)=0$$

KKT

$$\frac{\partial L}{\partial P_i} = \frac{df_i}{dP_i} - \lambda - \underline{\mu_i} + \overline{\mu_i} = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{N_g} P_i - C_g = 0$$

$$\frac{\partial L}{\partial \underline{\mu_i}} = P_i - \underline{P_i} \geq 0$$

$$\frac{\partial L}{\partial \overline{\mu_i}} = \overline{P_i} - P_i \geq 0$$

$$\underline{\mu_i}\left(P_i - \underline{P_i}\right) = 0$$

$$\overline{\mu_i}\left(\overline{P_i} - P_i\right) = 0$$

Comments

- In the solution point, if any generating group reaches its minimum or maximum limit, me marginal production cost is the same for all them:

$$\lambda = \frac{df_1}{dP_1} = \frac{df_2}{dP_2} = \ldots = \frac{df_{Ng}}{dP_{Ng}}$$

In this condition, the Lagrange multipliers associated to the non equality constraints are equal to zero

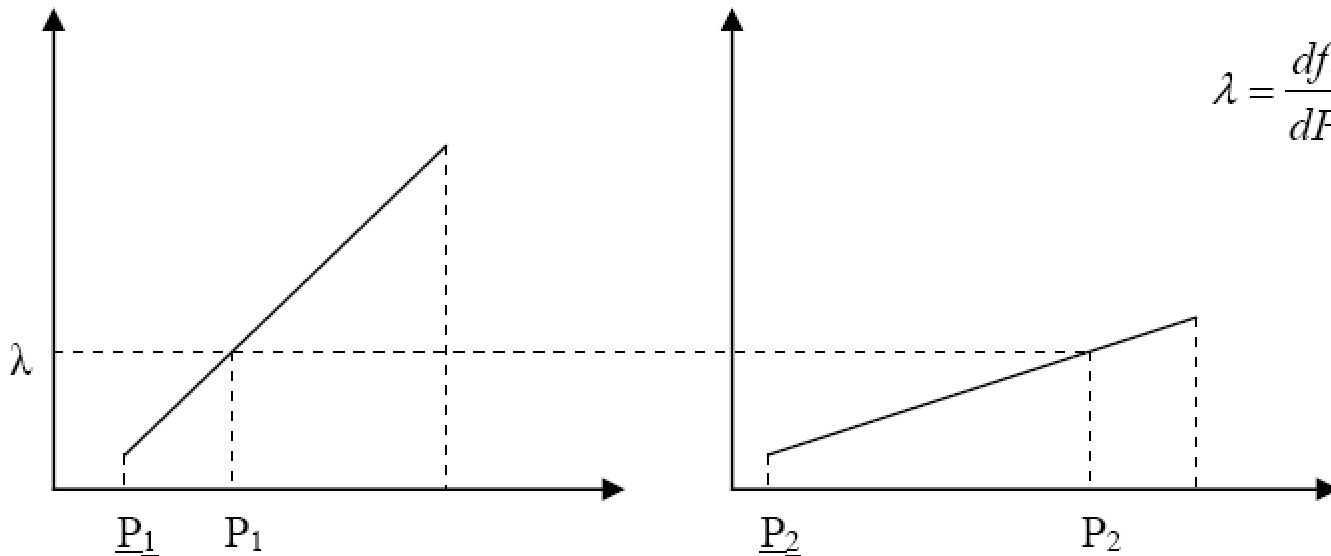
$$\underline{P_i} \leq P_i \leq \overline{P_i} \Rightarrow \overline{\mu_i} = \underline{\mu_i} = 0$$

If a unit reaches its minimum level $P_i = \underline{P_i} \Rightarrow \underline{\mu_i} \geq 0$

If a unit reaches its maximum level $P_i = \overline{P_i} \Rightarrow \overline{\mu_i} \geq 0$

Graphical method to find the solution

$$\frac{\partial L}{\partial P_1} = A_{11} + 2A_{21}P_1 - \lambda - \underline{\mu_1} + \overline{\mu_1} = 0$$
$$\frac{\partial L}{\partial P_2} = A_{12} + 2A_{22}P_2 - \lambda - \underline{\mu_2} + \overline{\mu_2} = 0$$
$$\frac{\partial L}{\partial \lambda} = P_1 + P_2 - C_g = 0$$
$$\frac{\partial L}{\partial \underline{\mu_1}} = P_1 - \underline{P_1} \geq 0$$
$$\frac{\partial L}{\partial \overline{\mu_1}} = \overline{P_1} - P_1 \geq 0$$
$$\frac{\partial L}{\partial \underline{\mu_2}} = P_2 - \underline{P_2} \geq 0$$
$$\frac{\partial L}{\partial \overline{\mu_2}} = \overline{P_2} - P_2 \geq 0$$
$$\underline{\mu_1}\left(P_1 - \underline{P_1}\right) = 0$$
$$\overline{\mu_1}\left(\overline{P_1} - P_1\right) = 0$$
$$\underline{\mu_2}\left(P_2 - \underline{P_2}\right) = 0$$
$$\overline{\mu_2}\left(\overline{P_2} - P_2\right) = 0$$



$$\lambda = \frac{df_1}{dP_1} = \frac{df_2}{dP_2}$$

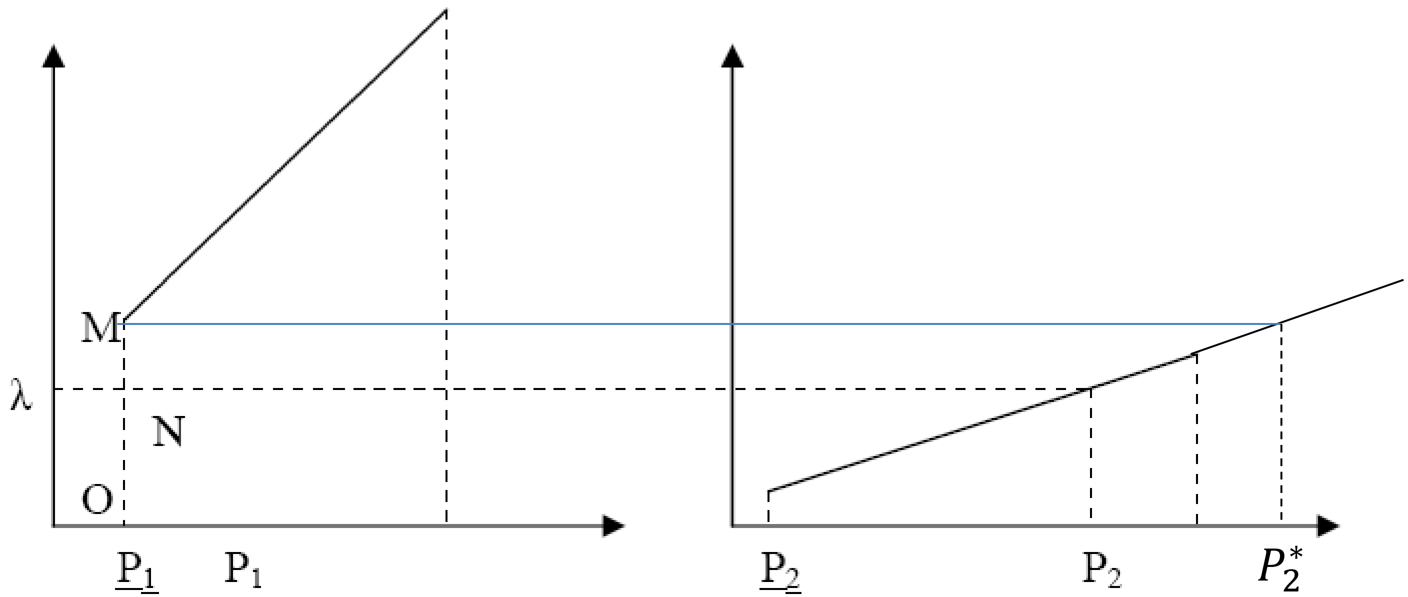
$P_1 + P_2 = C_g = \text{Load}$

Case A: no upper and lower bound reached

$$\overline{\mu_1} = \underline{\mu_1} = 0$$

$$\overline{\mu_2} = \underline{\mu_2} = 0$$

Unit 1 reaches its lower bound



Exclusion condition for unit 1:

$$\overline{\mu_1} = 0$$

$$\underline{\mu_1} > 0$$

$$\overline{\mu_2} = \underline{\mu_2} = 0$$

How to find the solution

$$\begin{cases} C_{11} + 2C_{21}P_1^* - \lambda - \underline{\mu_1} = 0 \\ C_{12} + 2C_{22}P_2^* - \lambda = 0 \end{cases}$$

From the second equation we obtain the Lagrange multiplier λ

$$\lambda = C_{12} + 2C_{22}P_2^*$$

And

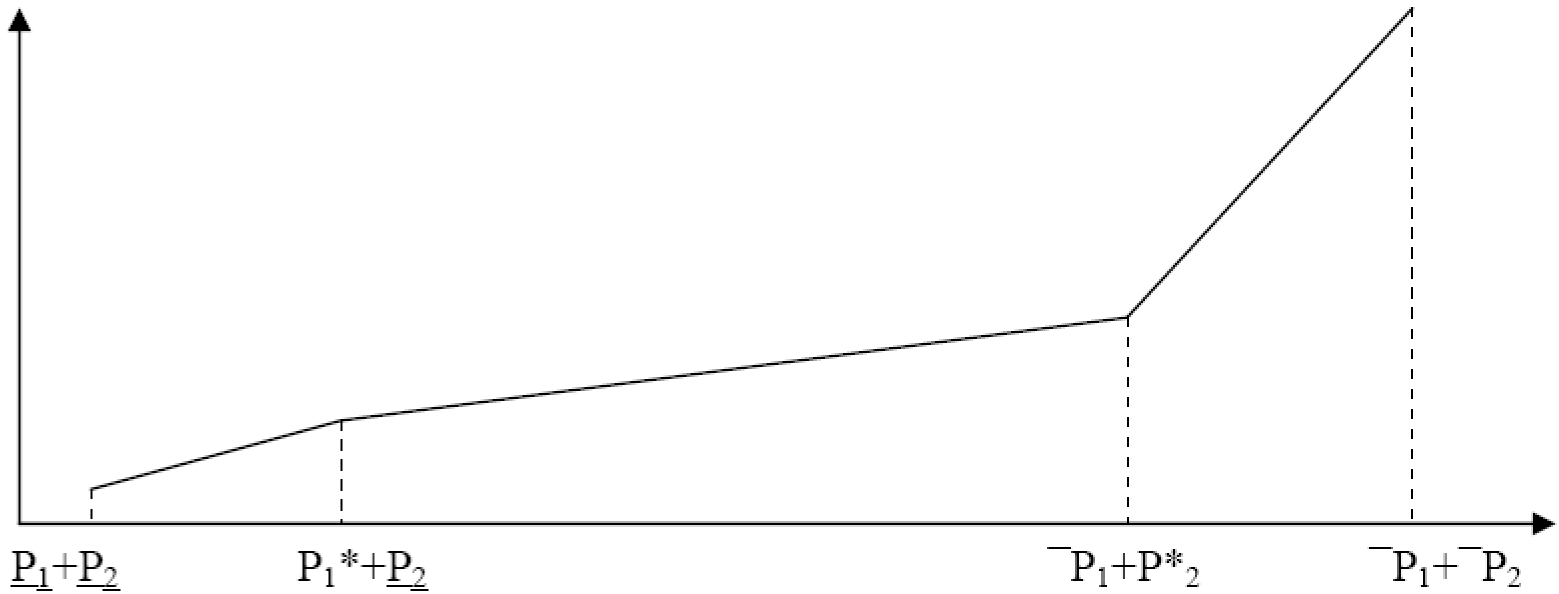
$\underline{\mu_1} = C_{11} + 2C_{21}P_1^* - \lambda = C_{11} + 2C_{21}P_1^* - C_{12} - 2C_{22}P_2^*$ =MO-NO>0

The Lagrange multiplier $\underline{\mu_1}$ is maximum when the power of the Unit 2 is minimum.

$\underline{\mu_1}$ decreases when the marginal production cost of Unit 2 increases.

$\underline{\mu_1}$ is zero when $P_2 = P_2^*$ and $P_1 = \underline{P_1}$

Equal marginal cost curve



This curve represents all solutions in the range $\underline{P_1} + \underline{P_2}$ and $\overline{P_1} + \overline{P_2}$

For each Load the marginal cost is reported on y-axis

The are 4 significant points in this curve

The β-dispatching problem

- The real losses on the network are considered
- The real losses are a function of the production profile
- In the equality constraint is introduced an expression that estimates the real losses

The model

Losses estimation function $$p = b_0 + \sum_{i=1}^{Ng-1} b_{1i} \cdot P_i + \frac{1}{2} \cdot \sum_{i=1}^{Ng-1} \sum_{j=1}^{Ng-1} b_{2ij} \cdot P_i \cdot P_j$$

$$\sum_{i=1}^{Ng} P_i = C_T + b_0 + \sum_{i=1}^{Ng-1} b_{1i} \cdot P_i + \frac{1}{2} \cdot \sum_{i=1}^{Ng-1} \sum_{j=1}^{Ng-1} b_{2ij} \cdot P_i \cdot P_j$$

The equality constraint is a non linear function

$$\ell = \sum_{i=1}^{Ng} f_i(P_i) - \lambda \left(\sum_{i=1}^{Ng} P_i - C_T - p\left(P_{i,i \neq ng}\right) \right) - \sum_{i=1}^{Ng} \underline{\pi_i} \cdot \left(P_i - \underline{P_i}\right) - \sum_{i=1}^{Ng} \overline{\pi_i} \cdot \left(\overline{P_i} - P_i\right)$$

The KKT conditions of the problem

$$\frac{\partial \ell}{\partial P_i} = \frac{df_i}{dP_i} - \lambda \cdot \left(1 - \frac{\partial p}{\partial P_i}\right) - \underline{\pi_i} + \overline{\pi_i} = 0$$

$$\beta_i = 1 - \frac{\partial p}{\partial P_i}$$ Penalty factor

$$\frac{\partial \ell}{\partial P_i} = \frac{df_i}{dP_i} - \lambda \cdot \beta_i - \underline{\pi_i} + \overline{\pi_i} = 0$$

$$\gamma_i = \begin{cases} 1 - b_{1i} - \dfrac{1}{2}\displaystyle\sum_{j=1}^{Ng-1} b_{2ij} \cdot P_j & i \neq ng \\ 1 & i = ng \end{cases}$$

$$\beta_i = \begin{cases} 1 - b_{1i} - \displaystyle\sum_{j=1}^{Ng-1} b_{2ij} \cdot P_j & i \neq ng \\ 1 & i = ng \end{cases}$$

$$\frac{\partial \ell}{\partial P_i} = \frac{df_i}{dP_i} - \lambda \cdot \left(1 - \frac{\partial p}{\partial P_i}\right) - \underline{\pi_i} + \overline{\pi_i} = \left(A_{1i} + 2A_{2i} \cdot P_i\right) - \lambda \cdot \beta_i - \underline{\pi_i} + \overline{\pi_i} = 0$$

$$\Rightarrow \qquad \frac{A_{1i}}{\beta_i} + \frac{2A_{2i}}{\beta_i} \cdot P_i - \lambda - \frac{\underline{\pi_i}}{\beta_i} + \frac{\overline{\pi_i}}{\beta_i} = 0$$

$$\sum_{i=1}^{Ng} P_i = \sum_{i=1}^{Ng-1} P_i + P_{ng} = C_T + p = C_T + b_0 + \sum_{i=1}^{Ng-1} b_{1i} P_i + \frac{1}{2} \sum_{i=1}^{Ng-1} \sum_{j=1}^{Ng-1} b_{2ij} P_i P_j = C_T + b_0 + \sum_{i=1}^{Ng-1} P_i \cdot \left(b_{1i} + \frac{1}{2} \sum_{j=1}^{Ng-1} b_{2ij} P_j \right) =$$

$$= C_T + b_0 + \sum_{i=1}^{Ng-1} P_i \cdot (1 - \gamma_i) = C_T + b_0 + \sum_{i=1}^{Ng-1} P_i - \sum_{i=1}^{Ng-1} \gamma_i P_i$$

$$\Rightarrow \qquad P_{ng} + \sum_{i=1}^{Ng-1} \gamma_i P_i = C_T + b_0$$

KKT conditions

$$\boxed{\begin{aligned}
&\frac{A_{1i}}{\beta_i} + \frac{2A_{2i}}{\beta_i} \cdot P_i - \lambda - \frac{\underline{\pi_i}}{\beta_i} + \frac{\overline{\pi_i}}{\beta_i} = 0 \\
&P_{ng} + \sum_{i=1}^{Ng-1} \gamma_i P_i = C_T + b_0 \\
&\underline{P_i} \le P_i \le \overline{P_i} \\
&\underline{\pi_i} \cdot \left(P_i - \underline{P_i} \right) = 0 \\
&\overline{\pi_i} \cdot \left(\overline{P_i} - P_i \right) = 0
\end{aligned}}$$

It is possible to define the following new variables:

$$P_i^* = \gamma_i P_i, \quad \underline{\pi_i}^* = \frac{\underline{\pi_i}}{\beta_i} \text{ e } \quad \overline{\pi_i}^* = \frac{\overline{\pi_i}}{\beta_i}$$

And write the KKT conditions as a function of them: to find the solution an interactive procedure is necessary where at each step a classical λ-dispatching problem in solved and then the coefficients β_i and γ_i are update

$$\frac{A_{1i}}{\beta_i} + \frac{2A_{2i}}{\beta_i \gamma_i} \cdot P_i^* - \lambda - \underline{\pi}_i^* + \overline{\pi}_i^* = 0$$

$$P_{ng} + \sum_{i=1}^{Ng-1} P_i^* = C_T + b_0$$

$$\underline{P_i}^* \le P_i^* \le \overline{P_i}^*$$

$$\underline{\pi_i}^* \cdot \left(P_i^* - \underline{P_i}^*\right) = 0$$

$$\overline{\pi_i}^* \cdot \left(\overline{P_i}^* - P_i^*\right) = 0$$

Algorithm

- Step 1: At iteration 0, a classical λ-dispatching problem is solved with an initial constant estimation of losses

- Step 2: The coefficients β_i and γ_i are update (they are a function of the production profile). The new variables P^* are introduced in the problem and a new λ-dispatching problem is defined

- Step 3: The new λ-dispatching problem is solved and in the solution point the original variable are computed. If there is a small variation of the OF value and the variation of the control variables is lower than a tolerance, stop, otherwise go to step 2

Comments

- If no unit reaches its upper or lower bound, the KKT in the optimal point are the following:

$$\frac{df_i}{dP_i} = \lambda \cdot \left(1 - \frac{\partial \ell}{\partial P_i}\right) = \lambda \beta_i \qquad \Rightarrow \qquad \lambda = \frac{1}{\beta_i} \frac{df_i}{dP_i}$$

- For the slack bus:

$$\frac{\partial p}{\partial P_{ng}} = 0 \qquad \Rightarrow \qquad \frac{df_{Ng}}{dP_{Ng}} = \lambda$$

Comments

- The marginal production cost is not the same for all generating units because their impact on the losses is different:
  - With a fixed load, the total production (and therefore the cost production) changes if we have a different production profile because we have different losses
- If a units reaches its upper or lower bounds, the π multiplier will be different from zero
- Generally, $\beta i \simeq 1$
- No security network constraints are considered in the model

β_i estimation

$$\sum_{i=1}^{Ng} P_i = C_T + p \quad \Rightarrow \quad P_{ng} + \sum_{i=1}^{Ng-1} P_i = C_T + p\,.$$

We compute the derivative of the equality constraint with respect to the real power of the power unit i

$$\frac{\partial P_{ng}}{\partial P_i} + 1 = \frac{\partial p}{\partial P_i}$$

$$\beta_i = 1 - \frac{\partial p}{\partial P_i} = -\frac{\partial P_{ng}}{\partial P_i}$$

Example

- We have two power plants A and B and we assume that B is the slack bus.
- If I increase the production of Unit A for 1 MW, and if the losses are equal to 0, and the load is fixed, the Unit B have to reduce its output of 1 MW.
- If I increase the production of Unit A for 1 MW, and if the losses are considered, the Unit B have to reduce its output of a quantity different from 1 (higher or lower than 1 MW).
- In this way it is possible to estimate β_i coefficients

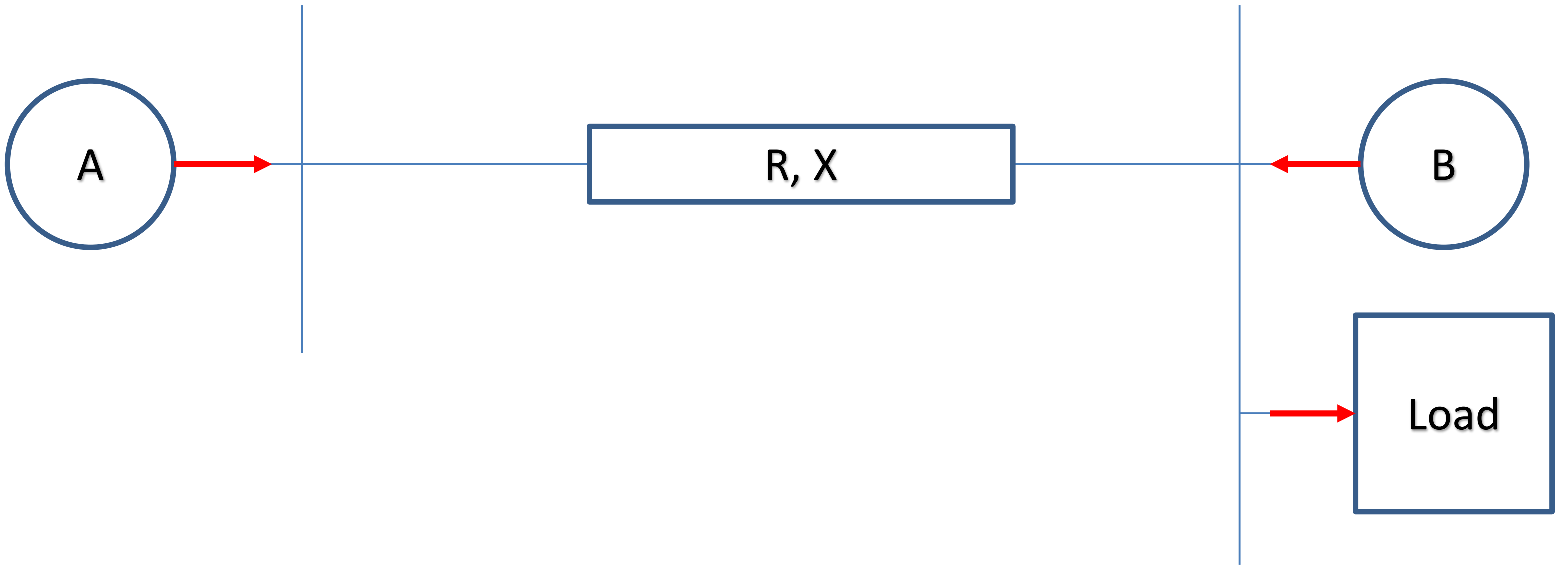
Example

- If the losses decrease, the reduction of the Unit B is higher than 1 MW:

$$\frac{\partial P_{ng}}{\partial P_1} = -\beta_1 \leq -1$$

- If the losses increase, the reduction of the Unit B is lower than 1 MW:

$$\frac{\partial P_{ng}}{\partial P_1} = -\beta_1 \geq -1$$



In this example, if the production of unit A is increased, we have a increment of the losses. Vice versa, if we increase the production of unit B, there is a reduction of the losses

If the total power required by the Load is produced by B, the losses are zero