DUAL METHODS

Pierre DUYSINX
LTAS – Automotive Engineering
University of Liege

Lagrange Function: an introduction to the concept

- Optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0 \qquad j = 1, \dots, m \end{aligned}$$

Where functions f(x) and g_j(x) are *continuous and differentiable*

- Lagrange function

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x})$$

With $\lambda_j \geq 0$ the Lagrange multipliers associated to each constraint

- Quasi-unconstrained problem called primal problem

$$\min_{\mathbf{x}} \max_{\lambda \geq 0} \quad L(\mathbf{x}, \lambda)$$

Karush-Kuhn-Tucker optimality conditions

- NECESSARY OPTIMALITY CONDITIONS

  If x^* is a regular point (i.e. gradient of active constraints are linearly independent in x^*)
  if $f(x)$ and $g_j(x)$ are continuous and differentiable

  Then a necessary condition for x^* to be an optimum of the optimization problem is to be able to find a set of Lagrange Multipliers λ^* such that:

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j(\mathbf{x}^*)}{\partial x_i} = 0 \quad \forall\, i$$

$$g_j(\mathbf{x}^*) \leq 0$$

$$\lambda_j^* \geq 0$$

$$\lambda_j^* \, g_j(\mathbf{x}^*) = 0 \quad \forall\, j$$

Introduction to Lagrangian duality

- KKT optimality conditions

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j(\mathbf{x}^*)}{\partial x_i} = 0 \qquad \forall\ i$$

  are equivalent to optimality conditions of minimization problem

$$\min_{\mathbf{x}} \; L(\mathbf{x}, \lambda^\star)$$

- Imagine now that for any $\lambda>0$, we can solve the minimization problem

$$\min_{\mathbf{x}} \; L(\mathbf{x}, \lambda)$$

  This minimization problem is called Lagrangian problem

- The solution of Lagrangian problem establishes a *relation of dependency between primal and dual variables*

$$\mathbf{x} = \mathbf{x}(\lambda) = \arg \min_{\mathbf{x}} \; L(\mathbf{x}, \lambda)$$

Introduction to Lagrangian duality

- Inserting these relations into the Lagrange function gives rise to:

$$\begin{aligned} \ell(\lambda) &= L(\mathbf{x}(\lambda), \lambda) \\ &= f(\mathbf{x}(\lambda)) + \sum_{j=1}^{m} \lambda_j \, g_j(\mathbf{x}(\lambda)) \end{aligned}$$

where function $\ell(\lambda)$ is called dual function

- Optimization problem related to the dual function is a maximization problem, called dual problem

$$\begin{aligned} \max_{\lambda_j} \quad & \ell(\lambda) \\ \text{s.t.} \quad & \lambda_j \geq 0 \qquad j = 1, \dots, m \end{aligned}$$

Introduction to Lagrangian duality

- **Unsolved question in this introduction**: is it possible (allowed) to swap min max operations ?

Primal Problem

$$\min_{\mathbf{x}} \max_{\lambda \geq 0} \quad L(\mathbf{x}, \lambda)$$

Dual Problem

$$\begin{aligned} \max_{\lambda_j} \quad & \ell(\lambda) \\ \text{s.t.} \quad & \lambda_j \geq 0 \qquad j = 1, \dots, m \end{aligned}$$

where $\ell(\lambda) = \min_{\mathbf{x}} L(\mathbf{x}, \lambda)$

Weak duality

- With very little assumptions on f(x) and $g_j(x)$ (even for non convex or discrete valued-problems), it is always possible to define a dual function in the following (general) way:

$$\ell(\lambda) = \inf_{\mathbf{x}} \quad L(\mathbf{x}, \lambda)$$

- One can prove that:
  - The dual function is *concave*
  - The dual function may be *non-smooth*

- Dual problem

$$\begin{aligned} \max_{\lambda_j} \quad & \ell(\lambda) \\ \text{s.t.} \quad & \lambda_j \geq 0 \qquad j = 1, \dots, m \end{aligned}$$

Weak Duality

- **General properties of dual function**
  - Dual function is a *lower bound of primal objective function on the feasible set*

$$f(\mathbf{x}) \geq \ell(\lambda)$$

  - Optimal value of dual function is a lower bound of optimal value of objective function

$$f(\mathbf{x}^\star) \geq \ell(\lambda^\star)$$

  - Duality gap

$$G = f(\mathbf{x}^\star) - \ell(\lambda^\star)$$

Strong Duality

- Convex problem

$$\begin{aligned} \min_{\mathbf{x} \in X} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & g_j(\mathbf{x}) \leq 0 \qquad j = 1, \ldots, m \end{aligned}$$

with f(x), g_j(x) which are C^1 and *convex*
X a convex set, for example (side-constraints)

$$X = \{x_i \mid \underline{x}_i \leq x_i \leq \overline{x}_i \quad i = 1 \ldots n\}$$

- Slater condition is satisfied

$$\exists\, \tilde{\mathbf{x}} : g_j(\tilde{\mathbf{x}}) < 0 \;\forall j$$

Strong Duality

- If the problem is convex and if Slater condition is satisfied,
Then there is at least one Lagrange multiplier vector λ* such that

$$f(\mathbf{x}^\star) = \min_{\mathbf{x}} L(\mathbf{x}, \lambda^\star)$$

and there is no duality gap

$$f(\mathbf{x}^\star) = \ell(\lambda^\star)$$

- Corollary: x* is the optimal primal solution and λ* is the optimal dual solution, if and only if Karush-Kuhn-Tucker conditions are satisfied at (x*, λ*).

In other words, *KKT conditions are necessary and sufficient for convex problems* which satisfies Slater conditions.

Strong Duality

- Corollary: *Solving primal problem is totally equivalent to solving dual problem*:

$$\begin{aligned} \min_{\lambda_j} \quad & \ell(\lambda) \\ s.t. \quad & \lambda_j \geq 0 \qquad j = 1, \ldots, m \end{aligned}$$

with dual function

$$\ell(\lambda) = \min_{\mathbf{x} \in X} \quad L(\mathbf{x}, \lambda)$$

Indeed Lagrange function $L(\mathbf{x}, \lambda)$ has a saddle point in $(\mathbf{x}^\star, \lambda^\star)$, which means that

$$L(\mathbf{x}^\star, \lambda) \leq L(\mathbf{x}^\star, \lambda^\star) \leq L(\mathbf{x}, \lambda^\star)$$

so that

$$\min_{\mathbf{x} \in X} \max_{\lambda \geq 0} L(\mathbf{x}, \lambda) \quad \Leftrightarrow \quad \max_{\lambda \geq 0} \min_{\mathbf{x} \in X} L(\mathbf{x}, \lambda)$$

Dual function properties in strong duality

- Concavity
- No duality gap $\quad \ell(\lambda^\star) = f(\mathbf{x}^\star)$
- Gradient of dual function

$$\frac{\partial \ell(\lambda)}{\partial \lambda_k} = g_k(\mathbf{x}(\lambda))$$

Proof

$$\begin{aligned} \frac{\partial \ell(\lambda)}{\partial \lambda_k} &= \frac{\partial}{\partial \lambda_k} \{ f(\mathbf{x}(\lambda)) + \sum_{j=1}^{m} \lambda_j \, g_j(\mathbf{x}(\lambda)) \} \\ &= \sum_{i=1}^{n} \left[\frac{\partial f(\mathbf{x}(\lambda))}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(\mathbf{x}(\lambda))}{\partial x_i} \right] \frac{\partial x_i}{\partial \lambda_k} + g_k(\mathbf{x}(\lambda)) \end{aligned}$$

Dual function properties in strong duality

- Optimality conditions of quasi-unconstrained dual problem:

$$\frac{\partial \ell(\lambda)}{\partial \lambda_k} = 0 \quad \text{if} \quad \lambda_k > 0$$
$$\frac{\partial \ell(\lambda)}{\partial \lambda_k} < 0 \quad \text{if} \quad \lambda_k = 0$$

In optimum of dual space: constraints are active (g_j=0) when λ_j >0, whereas constraints are strictly satisfied (g_j<0) when λ_j =0.

Dual function properties in strong duality

- A first dual maximization algorithm (based on steepest <u>ascent</u> method)

$$\lambda^+ = \lambda + \alpha \nabla \ell$$
$$\lambda_j^+ = \lambda_j + \alpha\, g_j(\mathbf{x}(\lambda))$$

Interpretation of dual approach: dual method tries to satisfy the constraints by adjusting the value of the Lagrange multipliers

Elementary dual optimizer can be realized by modifying an unconstrained algorithms in order to take care of non negativity restrictions

Dual function properties in strong duality

- Separable case
  - Solving Lagrangian problem a lot of times can be expensive

$$\mathbf{x}(\lambda) = \arg \min_{\mathbf{x} \in X} L(\mathbf{x}, \lambda)$$

    except if the problem is separable !

  - Function f(x) is separable if $f(\mathbf{x}) = \sum_{i=1}^{n} f_i(x_i)$
  - Problem is separable if the objective function and every constraint functions are separable. In this case Lagrange function is separable too, and Lagrangian problems can be split into n one-dimensional problems, which often have a simple algebraic structure and can be solved in closed form

$$\min_{\mathbf{x} \in X} L(\mathbf{x}, \lambda) = \sum_{i=1}^{n} \min_{x_i \in X_i} L_i(x_i, \lambda)$$

Dual function properties in strong duality

- Second derivatives

$$\frac{\partial^2 \ell(\lambda)}{\partial \lambda_k \partial \lambda_l} = \frac{\partial}{\partial \lambda_l} g_k(\mathbf{x}(\lambda)) = \sum_{i=1}^{n} \frac{\partial g_k}{\partial x_i} \frac{\partial x_i(\lambda)}{\partial \lambda_l}$$

Evaluation of the derivatives of primal-dual relationships (from KKT conditions)

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j(\mathbf{x}^*)}{\partial x_i} = 0$$

with respect to λ_l:

$$\sum_{k=1}^{n} \left[\frac{\partial^2 f}{\partial x_i \partial x_k} + \sum_{j=1}^{m} \lambda_j \frac{\partial^2 g_j}{\partial x_i \partial x_k} \right] \frac{\partial x_k(\lambda)}{\partial \lambda_l} + \frac{\partial g_l(\lambda)}{\partial x_i} = 0$$

Dual function properties in strong duality

- Second order derivatives (cont'd)

$$\nabla^2_{xx} L(\mathbf{x}, \lambda) \frac{\partial \mathbf{x}(\lambda)}{\partial \lambda_l} + \nabla_x g_l = 0$$

must be restricted to free variables only.

Let be **G** the Hessian of Lagrangian matrix when rows and columns corresponding to fixed design variables have been deleted.

$$\mathbf{G} \frac{\partial \mathbf{x}(\lambda)}{\partial \lambda_l} + \nabla_x g_l = 0$$

Let **N**, the matrix whose columns are made of the constraint gradients

One gets:

$$\left[\frac{\partial^2 \ell(\lambda)}{\partial \lambda_k \partial \lambda_l} \right] = - \mathbf{N}^T \mathbf{G}^{-1} \mathbf{N}$$

Dual function properties in strong duality

- Second order derivatives (cont'd)

  Second order derivatives of dual function are discontinuous each time that the set of free and fixed design variables is modified.

  So the discontinuity of Hessian matrix happens along hyperplanes of equation:

$$x_i(\lambda) = \underline{x}_i \quad \text{and} \quad x_i(\lambda) = \overline{x}_i$$

- Second order algorithm for dual function maximization

  One has to take care of the discontinuity of the Hessian.

  Newton or quasi-Newton methods can only be applied within sub-regions of the dual workspace. Updating the set of free design variables is a difficult issue.

Application: solution of quadratic separable problem

- Problem statement:

$$\min_{\mathbf{x}} \quad \frac{1}{2}\mathbf{x}^T\mathbf{x} \qquad \text{s.t.} \quad \mathbf{C}^T\mathbf{x} \geq \mathbf{d}$$

where **C** is n x m matrix of constraint gradients (**C**=**N**(x) constant)

- Lagrange function $L(\mathbf{x},\lambda) = \frac{1}{2}\mathbf{x}^T\mathbf{x} - \lambda^T(\mathbf{C}^T\mathbf{x} - \mathbf{d})$

- KKT conditions (optimality conditions)

$$\nabla_x L = \mathbf{x} - \mathbf{C}\lambda = 0$$
$$-\mathbf{C}^T\mathbf{x} + \mathbf{d} \leq 0$$
$$\lambda \geq 0$$
$$\lambda^T(\mathbf{C}^T\mathbf{x} - \mathbf{d}) = 0$$

Application: solution of quadratic separable problem

- Primal-dual relationships: come from first set of KKT conditions, that is the solution of the Lagrangian problem

$$\mathbf{x}(\lambda) = \mathbf{C}\lambda$$

This solution is *fully explicit* here

- Dual function

$$\ell(\lambda) = \frac{1}{2}\mathbf{x}(\lambda)^T\mathbf{x}(\lambda) - \lambda^T(\mathbf{C}^T\mathbf{x}(\lambda) - \mathbf{d})$$
$$= \frac{1}{2}\lambda^T\mathbf{C}^T\mathbf{C}\lambda - \lambda^T(\mathbf{C}^T\mathbf{C}\lambda - \mathbf{d})$$
$$= -\frac{1}{2}\lambda^T\mathbf{C}^T\mathbf{C}\lambda + \lambda^T\mathbf{d}$$

Application: solution of quadratic separable problem

- Gradient of dual function:

$$\begin{aligned}\nabla\ell(\lambda) &= -\mathbf{C}^T\mathbf{C}\lambda + \mathbf{d} \\ &= \mathbf{d} - \mathbf{C}^T\mathbf{x}(\lambda)\end{aligned}$$

- Hessian of dual function (negative semi-definite so concavity)

$$\nabla^2\ell(\lambda) = -\mathbf{C}^T\mathbf{C}$$

- Dual function of a quadratic problem is a quadratic function

$$\ell(\lambda) = -\frac{1}{2}\lambda^T\mathbf{A}\lambda + \lambda^T\mathbf{d}$$

with $\mathbf{A} = \mathbf{C}^T\,\mathbf{C}$

Application: solution of quadratic separable problem

- Equality constraints
  - optimality condition in dual space:

$$\nabla\ell(\lambda) = -\mathbf{A}\lambda + \mathbf{d} = 0$$

  - dual optimal solution:

$$\lambda^\star = \mathbf{A}^{-1}\,\mathbf{d}$$

  - primal solution, recovered from primal dual relations:

$$\mathbf{x}^\star = \mathbf{C}\,\lambda^\star$$

Application: solution of quadratic separable problem

- Inequality constraints
  - dual maximization problem

$$\max_{\lambda} \quad -\frac{1}{2}\lambda^T \mathbf{A}\lambda \; + \; \lambda^T \mathbf{d}$$
$$\text{s.t.} \quad \lambda_j \; \geq \; 0$$

  - primal solution recovered with primal-dual relations:

$$\mathbf{x}^{\star} \; = \; \mathbf{C}\,\lambda^{\star}$$

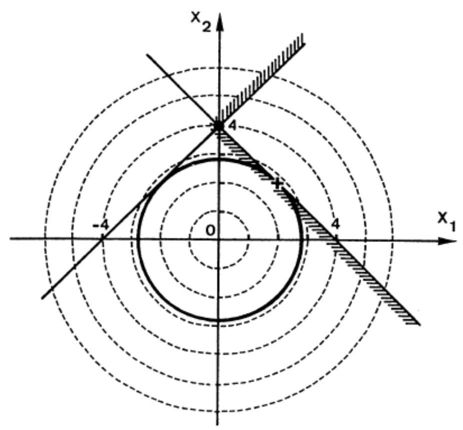
- Remarks:
  - dual problem is the most natural way to select the Lagrangian multipliers
  - dual problem is only a quasi-unconstrained problem
  - dimensionality of dual problem is generally lower than primal problem

Example of quadratic Problem

- Primal problem

$$\min_{x_1\,,\,x_2} \quad \frac{1}{2}\,x_1^2 \; + \; \frac{1}{2}\,x_2^2$$
$$\text{s.t.} \quad x_1 \; + \; x_2 \; \geq \; 4$$
$$x_1 \; - \; x_2 \; \geq \; -4$$

$$\mathbf{C} \; = \; \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \mathbf{d} \; = \; \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$



- Lagrange function

$$L(\mathbf{x}, \lambda) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 - \lambda_1(x_1 + x_2 - 4) - \lambda_2(x_1 - x_2 + 4)$$

- Lagrangian problem gives

$$x_1(\lambda_1, \lambda_2) = \lambda_1 + \lambda_2$$
$$x_2(\lambda_1, \lambda_2) = \lambda_1 - \lambda_2$$

Example of quadratic Problem

- Dual function: explicit !

$$\ell(\lambda_1, \lambda_2) = -(\lambda_1 - 2)^2 - (\lambda_2 + 2)^2 + 8$$

- Dual problem

$$\begin{aligned} \max_{\lambda_1, \lambda_2} \quad & \ell(\lambda_1, \lambda_2) = -(\lambda_1 - 2)^2 - (\lambda_2 + 2)^2 + 8 \\ \text{s.t.} \quad & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \end{aligned}$$

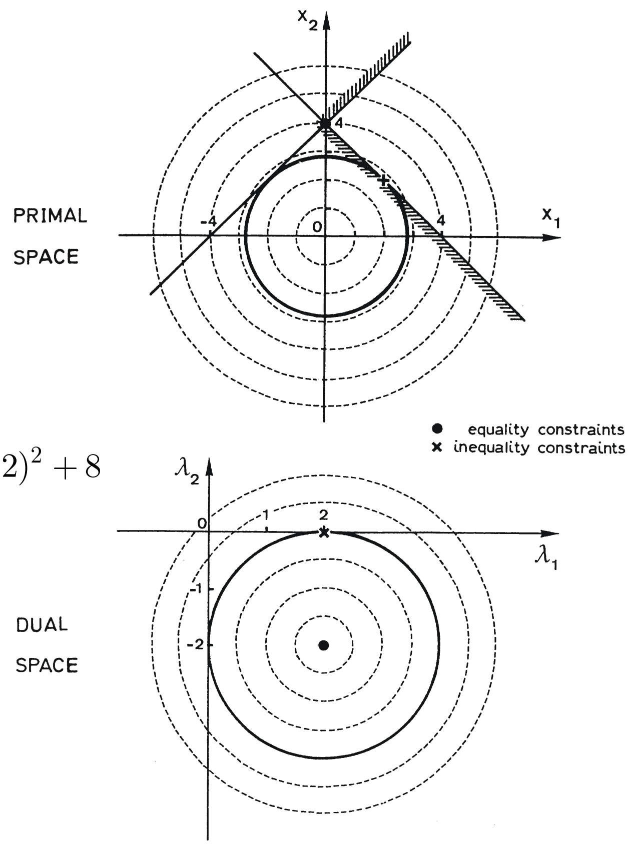
- Gradient of dual function = - Primal constraint values

$$\frac{\partial \ell}{\partial \lambda_1} = -2\lambda_1 + 4 = 4 - x_1 - x_2$$
$$\frac{\partial \ell}{\partial \lambda_2} = -2\lambda_2 - 4 = -4 - x_1 + x_2$$

Example of quadratic Problem

$$\begin{aligned} \min_{x_1, x_2} \quad & \frac{1}{2} x_1^2 + \frac{1}{2} x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 4 \\ & x_1 - x_2 \geq -4 \end{aligned}$$

$$\begin{aligned} \max_{\lambda_1, \lambda_2} \quad & \ell(\lambda_1, \lambda_2) = -(\lambda_1 - 2)^2 - (\lambda_2 + 2)^2 + 8 \\ \text{s.t.} \quad & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \end{aligned}$$



Treatment of side-constraints

- Quadratic separable problem with side constraints

$$\min_{\mathbf{x}} \quad \frac{1}{2}\sum_{i=1}^{n} x_i^2$$

$$\text{s.t.} \quad \sum_{i=1}^{n} c_{ij}\, x_i \;\geq\; d_j \quad j = 1 \ldots m$$

$$\underline{x}_i \;\leq\; x_i \;\leq \overline{x}_i \quad i = 1 \ldots n$$

- Advantage of separate treatment of side constraints:
  - if side constraints are treated as linear constraints, this increases dramatically the size of dual problem, and so the computational effort
  - The efficiency of dual problem solution calls for a separate treatment of these very simple structure constraints

Treatment of side-constraints

- Lagrangian problem

$$\min_{\underline{x}_i \;\leq\; x_i \;\leq \overline{x}_i} \quad \frac{1}{2}\sum_{i=1}^{n} x_i^2 \;-\; \sum_{j=1}^{m} \lambda_j \,(\sum_{i=1}^{n} c_{ij}\, x_i \;-\; d_j)$$

can be broken up into n one-dimensional Lagrangian problems

$$\min_{\underline{x}_i \;\leq\; x_i \;\leq \overline{x}_i} \quad L_i(x_i, \lambda) \;=\; \frac{1}{2}\, x_i^2 \;-\; (\sum_{j=1}^{m} \lambda_j\, c_{ij})\; x_i$$

The global solutions of these problems are given by expressing the optimality conditions

$$\frac{\partial L_i}{\partial x_i} \;=\; x_i - (\sum_{j=1}^{m} \lambda_j\, c_{ij}) \;=\; 0$$

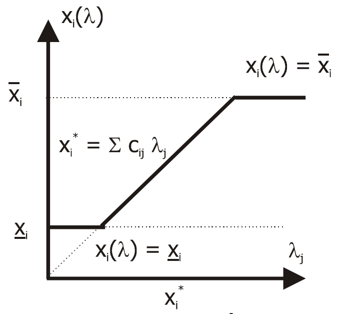
Treatment of side-constraints

- This yields the unconstrained solution of the problem

$$x_i^\star = \left(\sum_{j=1}^{m} \lambda_j \, c_{ij}\right)$$

- Primal-dual relationships are obtained by enforcing side-constraints

$$\begin{aligned} x_i(\lambda) &= x_i^\star \quad \text{if} \quad \underline{x}_i \le x_i^\star \le \overline{x}_i, \\ x_i(\lambda) &= \underline{x}_i \quad \text{if} \quad x_i^\star \le \underline{x}_i, \\ x_i(\lambda) &= \overline{x}_i \quad \text{if} \quad x_i^\star \ge \overline{x}_i. \end{aligned}$$



  - These relationships are still fully explicit, but contain conditions
  - The expressions are non smooth, the derivative being discontinuous when one free variable becomes fixed or conversely

Treatment of side-constraints

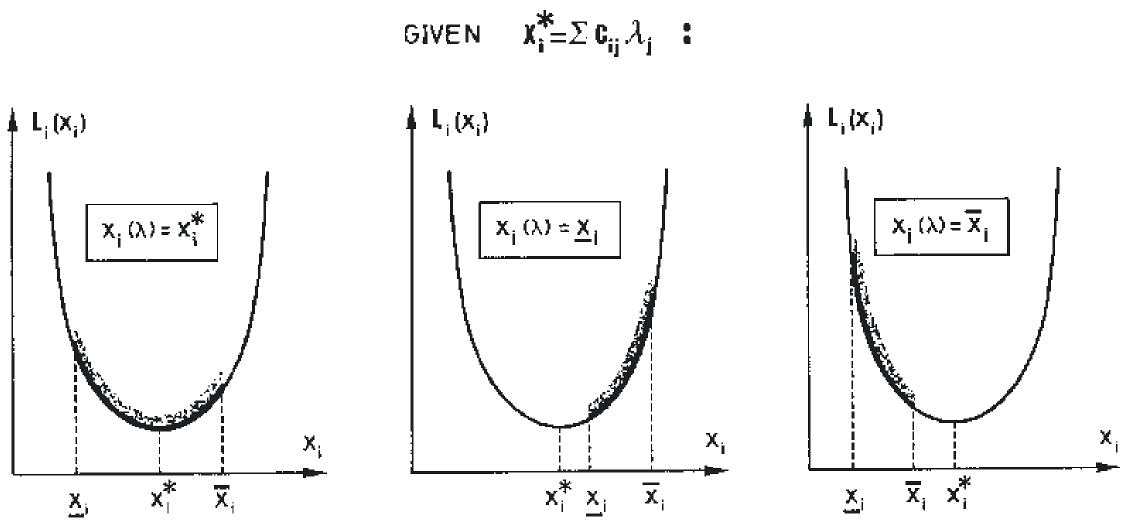


Illustration of the lagrangian problem with side-constraints

Treatment of side-constraints

- Dual function is obtained by inserting primal-dual relations into Lagrange function

$$\max_{\lambda} \quad \frac{1}{2}\sum_{i=1}^{n} x_i^2(\lambda) \; - \; \sum_{j=1}^{m} \lambda_j \, (\sum_{i=1}^{n} c_{ij}\, x_i(\lambda) \; - \; d_j)$$

$$\text{s.t.} \quad \lambda_j \; \geq \; 0$$

However, no explicit expression is available now !

- First derivatives are given by primal constraint values:

$$\frac{\partial \ell}{\partial \lambda_j} \; = \; d_j \; - \; (\sum_{i=1}^{n} c_{ij}\, x_i(\lambda))$$

Treatment of side-constraints

- Second derivatives:

$$A_{jk} \; = \; \frac{\partial^2 \ell}{\partial \lambda_j \partial \lambda_k} \; = \; \frac{\partial}{\partial \lambda_k}\left[d_j - \sum_{i=1}^{n} c_{ij}\, x_i(\lambda) \right] \; = \; \sum_{i=1}^{n} c_{ij} \, \frac{\partial x_i}{\partial \lambda_k}$$

Calculating the derivatives of the primal-dual relations, taking into account side-constraints and the discontinuous character of the relations:

For free variables: $$\frac{\partial x_i}{\partial \lambda_k} \; = \; c_{ik} \quad \text{if} \quad \underline{x}_i \; < \; x_i \; < \; \overline{x}_i$$

For fixed variables $$\frac{\partial x_i}{\partial \lambda_k} \; = \; 0 \quad \text{if} \quad x_i \; = \overline{x}_i \quad \text{or} \quad x_i \; = \; \underline{x}_i$$

Treatment of side-constraints

- Hessian matrix of dual function

$$A_{jk} = \frac{\partial^2 \ell}{\partial \lambda_j \partial \lambda_k} = = \sum_{i \in F} c_{ij}\, c_{ik}$$

where F is the set of free design variables

So the second order derivatives are discontinuous each time that the set of free design variable is modified, i.e. that a free design variable becomes fixed or vice-versa.

Dual space is separated into sub-regions

- Discontinuity planes

$$\left(\sum_{j=1}^{m} \lambda_j\, c_{ij}\right) = \underline{x}_i \qquad \left(\sum_{j=1}^{m} \lambda_j\, c_{ij}\right) = \overline{x}_i$$

Example of treatment of side-constraints

- Primal problem

$$\begin{aligned} \min_{x_1\,,\,x_2} \quad & x_1^2 + x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 4 \\ & x_1 - x_2 \geq -4 \\ & 1 \leq x_i \leq 4 \quad i = 1, 2 \end{aligned}$$

- Lagrangian problems

$$L(\mathbf{x}, \lambda) = x_1^2 + x_2^2 - \lambda_1\,(x_1 + x_2 - 4) - \lambda_2\,(x_1 - x_2 + 4)$$

- Stationary conditions: but side constraints?

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= x_1 - \lambda_1 - \lambda_2 \implies x_1 = \lambda_1 + \lambda_2 \\ \frac{\partial L}{\partial x_2} &= x_2 - \lambda_1 + \lambda_2 \implies x_2 = \lambda_1 - \lambda_2 \end{aligned}$$

Example of treatment of side-constraints

- One-dimensional minimization problems with side-constraints

$$\min_{1 \leq x_1 \leq 4} \quad L_1(x_1) = 1/2\, x_1^2 - \lambda_1\, x_1 - \lambda_2\, x_2$$

$$\min_{1 \leq x_2 \leq 4} \quad L_2(x_2) = 1/2\, x_2^2 - \lambda_1\, x_2 + \lambda_2\, x_2$$

- Admit primal-dual relations in closed-form $\mathbf{x} = \mathbf{x}(\lambda)$:

$$\begin{aligned} x_1 &= \lambda_1 + \lambda_2 \quad \text{if} \quad 1 \leq \lambda_1 + \lambda_2 \leq 4 \\ x_1 &= 1 \quad \text{if} \quad \lambda_1 + \lambda_2 \leq 1 \\ x_1 &= 4 \quad \text{if} \quad \lambda_1 + \lambda_2 \geq 4 \end{aligned}$$

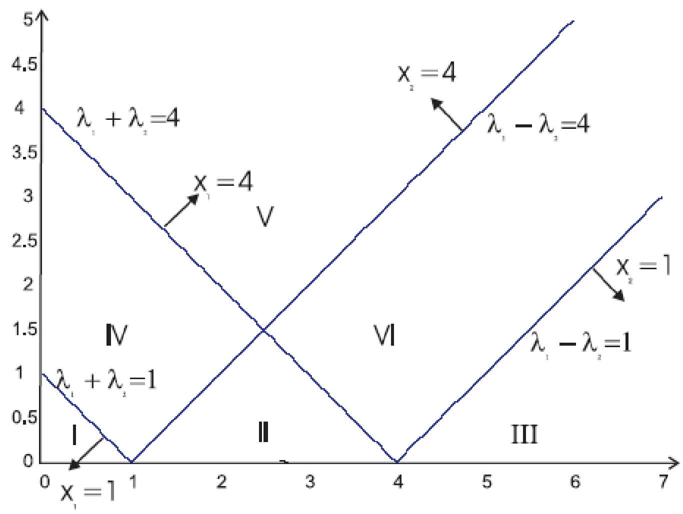
$$\begin{aligned} x_2 &= \lambda_1 - \lambda_2 \quad \text{if} \quad 1 \leq \lambda_1 - \lambda_2 \leq 4 \\ x_2 &= 1 \quad \text{if} \quad \lambda_1 - \lambda_2 \leq 1 \\ x_2 &= 4 \quad \text{if} \quad \lambda_1 - \lambda_2 \geq 4 \end{aligned}$$

Example of treatment of side-constraints

- Introduces 4 discontinuity planes

$$\begin{aligned} \lambda_1 + \lambda_2 &= 4 \\ \lambda_1 + \lambda_2 &= 1 \\ \lambda_1 - \lambda_2 &= 4 \\ \lambda_1 - \lambda_2 &= 1 \end{aligned}$$

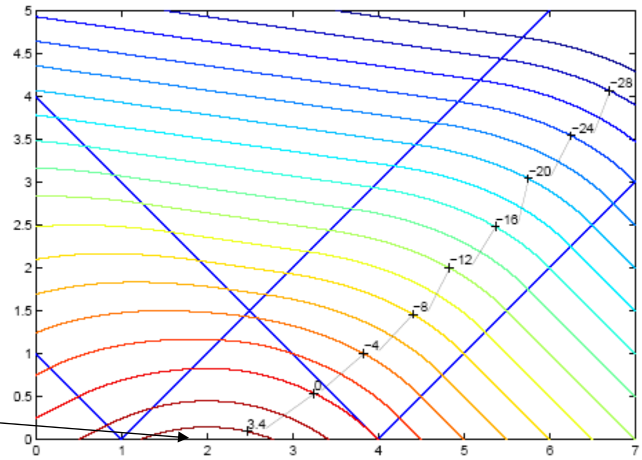
- And cut the dual space into 6 regions



Example of treatment of side-constraints

- Inserting the primal-dual relationships gives an explicit expression of the dual function in 6 regions

$\lambda_1^*=2$, $\lambda_2^*=0$



Region	x_1	x_2	$\ell(\lambda_1, \lambda_2)$
I	1	1	$1 + 2\lambda_1 - 4\lambda_2$
II	$\lambda_1 + \lambda_2$	$\lambda_1 - \lambda_2$	$-\lambda_1^2 - \lambda_2^2 + 4\lambda_1 - 4\lambda_2$
III	4	4	$16 - 4\lambda_1 - 4\lambda_2$
IV	$\lambda_1 + \lambda_2$	1	$-0.5\lambda_1^2 - 0.5\lambda_2^2 - \lambda_1\lambda_2 + 3\lambda_1 - 3\lambda_2 + 0.5$
V	4	1	$17/2 - \lambda_1 - 7\lambda_2$
VI	4	$\lambda_1 - \lambda_2$	$-0.5\lambda_1^2 - 0.5\lambda_2^2 + \lambda_1\lambda_2 - 8\lambda_2 + 8$

Dual method for MMA sub-problems

- MMA subproblems (after normalization):

$$\min_{\mathbf{x}} \quad \sum_{i=1}^{n} \frac{p_{i0}}{U_{ij} - x_i} + \sum_{i=1}^{n} \frac{q_{i0}}{x_i - L_{ij}}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} \frac{p_{ij}}{U_{ij} - x_i} + \sum_{i=1}^{n} \frac{q_{ij}}{x_i - L_{ij}} \le d_j \quad j = 1 \ldots m$$

$$\underline{x}_i \le x_i \le \overline{x}_i \quad i = 1 \ldots n$$

Remark: asymptotes U_{ij} and L_{ij} can depend on both variable index i and constraint index j

- Lagrange function (where we have introduced λ_0=1 for simplicity)

$$L(\mathbf{x}, \lambda) = \sum_{j=0}^{m} \lambda_j \left(\sum_{i=1}^{n} \frac{p_{ij}}{U_{ij} - x_i} + \sum_{i=1}^{n} \frac{q_{ij}}{x_i - L_{ij}} - d_j\right)$$

Dual method for MMA sub-problems

- The Lagrangian problem:

$$\min_{\underline{x}_i \le x_i \le \overline{x}_i} \quad L(\mathbf{x}, \lambda)$$

- Because of separability, the n-dimensional problem can be split into n 1-dimensional problems:

$$\min_{\underline{x}_i \le x_i \le \overline{x}_i} \quad L_i(x_i, \lambda) \;=\; \sum_{j=0}^{m} \frac{\lambda_j \, p_{ij}}{U_{ij} - x_i} + \sum_{j=0}^{m} \frac{\lambda_j \, q_{ij}}{x_i - L_{ij}}$$

- For pure MMA problems, asymptotes depend only on variable index j, i.e. $U_{ij} = U_j$ and $L_{ij} = L_j$, so that Lagrangian problem can be solved in closed form from optimality conditions:

$$\frac{\sum_{j=0}^{m} \lambda_j \, p_{ij}}{(U_i - x_i)^2} - \frac{\sum_{j=0}^{m} \lambda_j \, q_{ij}}{(x_i - L_i)^2} \;=\; 0$$

Dual method for MMA sub-problems

- Primal-dual relationships for pure MMA:

$$x_i^\star(\lambda) \;=\; \frac{U_i \,+\, \eta L_i}{\eta + 1} \qquad \text{with} \qquad \eta \;=\; \sqrt{\frac{\sum_{j=0}^{m} \lambda_j \, p_{ij}}{\sum_{j=0}^{m} \lambda_j \, q_{ij}}}$$

- For generalize MMA family, each constraint has its own set of asymptotes, and it is not possible anymore to find the closed form solution of Lagrangian problem. Solution is obtained by resorting to an iterative Newton scheme applied to optimality conditions

$$x_i(\lambda^+) \;=\; x_i(\lambda) \;-\; \frac{\partial L_i / \partial x_i}{\partial^2 L_i / \partial x_i^2}$$

with

$$\frac{\partial L_i}{\partial x_i} \;=\; \sum_{j=0}^{m} \frac{\lambda_j \, p_{ij}}{(U_{ij} - x_i)^2} \;-\; \sum_{j=0}^{m} \frac{\lambda_j \, q_{ij}}{(x_i - L_{ij})^2}$$

$$\frac{\partial^2 L_i}{\partial x_i^2} \;=\; 2 \sum_{j=0}^{m} \frac{\lambda_j \, p_{ij}}{(U_{ij} - x_i)^3} \;-\; 2 \sum_{j=0}^{m} \frac{\lambda_j \, q_{ij}}{(x_i - L_{ij})^3}$$

Dual method for MMA sub-problems

- Primal-dual relations with treatment of side-constraints:

$$
\begin{aligned}
x_i(\lambda) &= x_i^\star \quad \text{if} \quad \underline{x}_i \le x_i^\star \le \overline{x}_i, \\
x_i(\lambda) &= \underline{x}_i \quad \text{if} \quad x_i^\star \le \underline{x}_i, \\
x_i(\lambda) &= \overline{x}_i \quad \text{if} \quad x_i^\star \ge \overline{x}_i.
\end{aligned}
$$

- Dual function, gradient of dual function:
  - calculate x = x(λ)
  - compute f(x(λ)) and g(x(λ))
  - insert the calculated values in L(x,λ) = l(λ)
  - gradient is just given by g(x(λ)).
- If sub-problem is too complicated, sub-problem solution is itself decomposed into a sequence of quadratic separable sub-sub-problems that can be efficiently solved by dual method (see before).