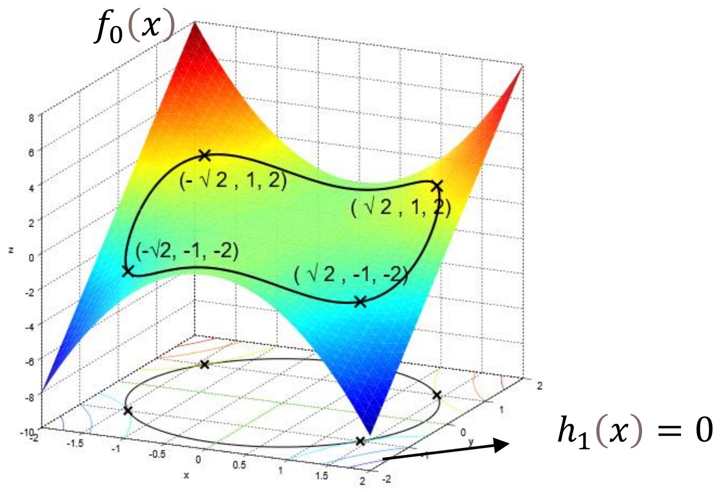
Linear Algebra & Convex Optimization – Lecture 13

Lagrange Method

Convex Optimization Formulation

General Optimization Formulation:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$



Convex Optimization Formulation:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & a_i^T x = b_i, \quad i = 1, \dots, p \end{array}$$

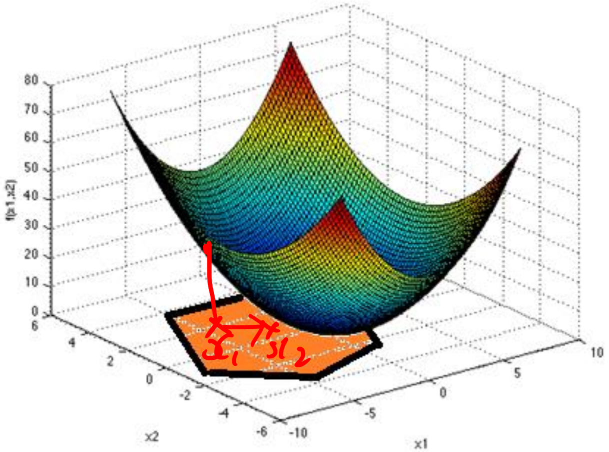
Poll: Suppose we consider $f_0(x)$ that is convex, will the above formulation convex ? Yes/No

$f_0, f_1, \dots, f_m$ are convex; equality constraints are affine

Feasible region of a convex optimization problem is convex

Alternative way of writing:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$



Convex formulation

Lagrangian Function

Solution $f_0(x) = p^*$

standard form problem (not necessarily convex):

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

$x \in \mathbf{R}^n$

$\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \;\cap\; \bigcap_{i=1}^{p} \mathbf{dom}\, h_i$

optimal value $p^\star$

$f_0(x) : R^n \to R$

Lagrangian: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\mathbf{dom}\, L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$,

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

$\lambda = [\lambda_1, \lambda_2 \ldots \lambda_m]$

$\nu = [\nu_1, \nu_2 \ldots \nu_p]$

weighted sum of objective and constraint functions

λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$

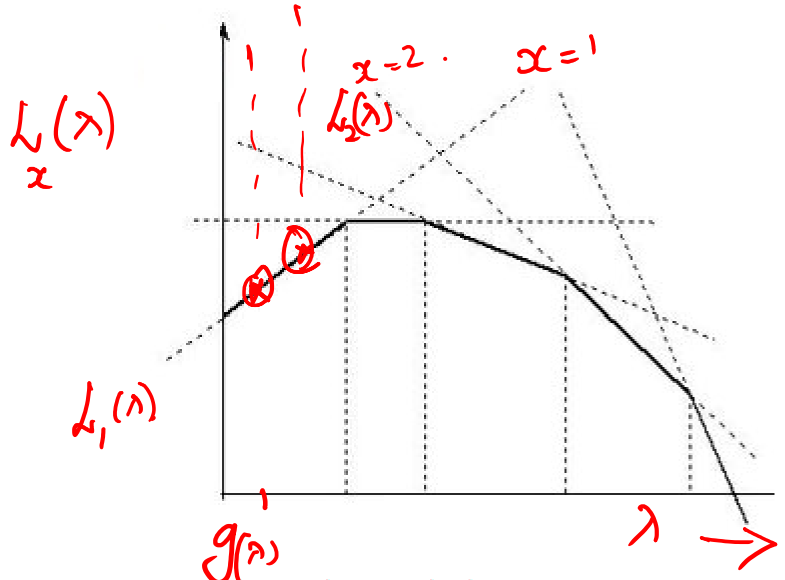
ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange Dual Function

Lagrange dual function: $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu)$$
$$= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

g is concave as it is point-wise infimum of family of affine functions in (λ, v)



Pointwise Minimum - Concave

Proof :

Consider a slight change of notation $L_x(\xi) = L(x, \lambda, v)$ where $\xi = (\lambda, v)$

For a fixed x , $L_x(\xi)$ is affine in ξ $\Rightarrow$ $\{L_x(\xi): x \in D\}$ is a family of affine functions

$g(\xi) = \inf_x \{ L_x(\xi): x \in D \}$ is point-wise infimum on a family of affine functions

Hence $g(\xi) = g(\lambda, v)$ is concave

Lagrange dual function in Lagrange variables is concave irrespective of the original functions

Lagrange Dual Function : Lower Bound Property

Standard Form Problem (optimal value p*):

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

primal problem

Lagrangian:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

Lagrange Dual Function:

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

$$\boxed{\text{if } \lambda \succeq 0, \text{ then } g(\lambda, \nu) \leq p^\star}$$

Proof :

$\lambda_i \geq 0$

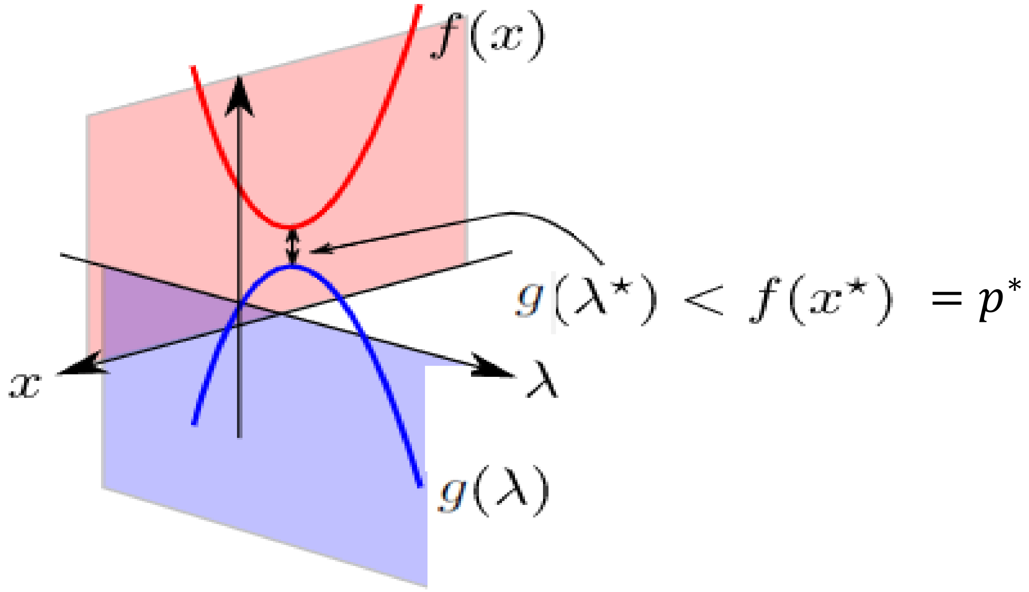
$$\tilde{x} \text{ is a feasible point} \Rightarrow \quad f_i(\tilde{x}) \leq 0 \quad h_i(\tilde{x}) = 0 \quad \Rightarrow \quad \sum_{i=1}^{m} \lambda_i f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x}) \leq 0,$$

$$L(\tilde{x}, \lambda, \nu) = f_0(\tilde{x}) + \underbrace{\sum_{i=1}^{m} \lambda_i f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x})}_{\leq 0} \leq f_0(\tilde{x})$$

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq L(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x})$$

Lagrange Dual Function : Illustration

Combined graph depiction of Dual and Original Problems

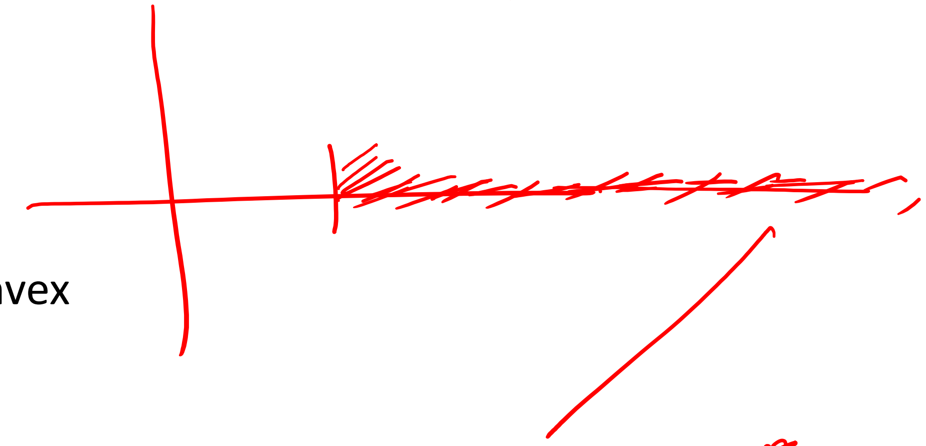


Lagrange Dual Optimization Problem

$\lambda \geq 0$

$\lambda \succeq 0$

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$



Dual problem is a convex optimization problem

Dual problem has a global maximum, d^* , even if original problem is non-convex

Dual problem find best lower bound on primal problem's optimal value, p^*

~~$x \succeq 0$~~

$x > 0$

λ, v are dual feasible if , $\lambda \geq 0$, $(\lambda, v) \in \mathbf{dom}\, g$

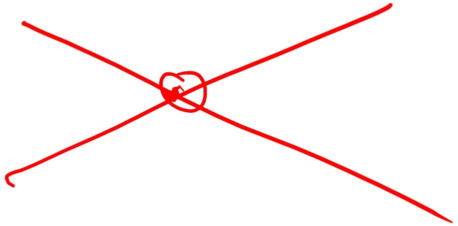
$x < 0$

Weak Duality:

$$d^\star \leq p^\star$$

$p^\star = -\infty \quad \Longrightarrow \quad d^\star = -\infty \quad \Longrightarrow$ Lagrangian dual problem is infeasible

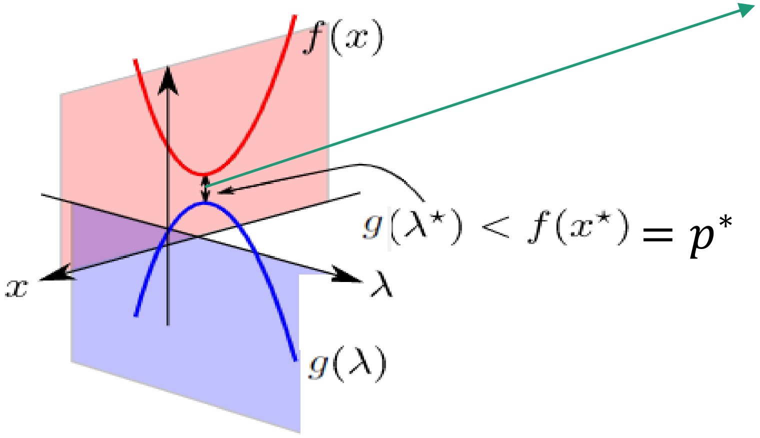
$d^\star = \infty \quad \Longrightarrow \quad p^\star = \infty \quad \Longrightarrow$ Primal problem is infeasible



Strong Duality

Strong duality : $d^{*} = p^{*}$

Does not hold in general ; **holds for convex problems**



Primal – Dual Gap

Primal & Dual Problems

Primal – Dual Gap is zero for strong duality

Complementary Slackness

$x \in \mathbb{R}^n$ $g(\lambda, \nu)$

Assume strong duality holds for the given problem ::

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \le 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

$x^\star$ is primal optimal $(\lambda^\star, \nu^\star)$ is dual optimal

$L(x, \lambda^\star, \nu^\star)$

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\cancel{\le}\ = \quad f_0(x^\star) + \underbrace{\sum_{i=1}^m \lambda_i^\star f_i(x^\star) + \sum_{i=1}^p \nu_i^\star h_i(x^\star)}_{=0} \\ &\cancel{\le}\ = \quad f_0(x^\star) \end{aligned}$$

The 2 inequalities can be replaced with equalities

This Implies :

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, \ldots, m$ (known as complementary slackness)

$\lambda_i^\star > 0 \Longrightarrow f_i(x^\star) = 0, \quad f_i(x^\star) < 0 \Longrightarrow \lambda_i^\star = 0$

Gradient of $L(x, \lambda^*, \nu^*)$ evaluated at x^* is zero, $\nabla L(x^*, \lambda^*, \nu^*) = 0$ ·

KKT (Karush-Kuhn-Tucker) Conditions

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

For Convex Problems: KKT conditions are sufficient for strong duality

Any pair of primal & dual points that satisfy KKT conditions are optimal points and have zero duality gap

$$\begin{array}{rcll} f_i(x^\star) & \leq & 0, & i = 1, \ldots, m \\ h_i(x^\star) & = & 0, & i = 1, \ldots, p \\ \lambda_i^\star & \geq & 0, & i = 1, \ldots, m \\ \lambda_i^\star f_i(x^\star) & = & 0, & i = 1, \ldots, m \end{array}$$

$$\nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(x^\star) = 0.$$

In a few special cases, KKT conditions can be solved analytically to obtain the solution.

Lagrange Method : Least Norm

Least Norm solution with Constraints:

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array} \qquad A \in \mathbf{R}^{p\times n}$$

$$\{\, x \mid Ax = b \,\} = \{\, x_p + z \mid z \in \mathcal{N}(A) \,\}$$

Lagrangian : $L(x, \nu) = x^T x + \nu^T (Ax - b)$ Domain: $\mathbf{R}^n \times \mathbf{R}^p$

KKT conditions:

$\nabla_x L(x, v^*) = 2x + A^T v^*$

$\nabla_x L(x, v^*) = 0$ at $x = x^*$

$2x^* + A^T v^* = 0 \implies x^* = (-1/2) A^T v^*$

$$\frac{\partial}{\partial x} L(x, v^*)$$

$$= \frac{\partial}{\partial x}\left[x^T x + \underbrace{v^{*T}(Ax - b)}\right]$$

$$= 2x + A^T v^*$$

$$\frac{\partial \mathbf{x}^\top \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{A} + \mathbf{A}^\top)\,\mathbf{x}$$

$$\frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^\top$$

Solution :

Solve for $Ax^* = b$ & $x^* = (-1/2) A^T v^*$ together

$-1/2\, AA^T v^* = b \qquad v^* = -2(AA^T)^{-1} b$

This gives, $\boxed{x^* = A^T (AA^T)^{-1} b}$

$x = (A^T A)^{-1} A^T b$ is the least square solution (For Tall A)

Lagrange Method : Constrained Least Squares

$[\; A \;]$

Least Squares solution of linear equations:

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|^2 \\ \text{subject to} & Cx = d. \end{array} \quad \text{— ②}$$

$$\frac{\partial}{\partial x}(L x, v^*) = \frac{\partial}{\partial x}\left((x^T A^T - b^T)(Ax - b) + v^{*T}(Cx)\right)$$

$$= \frac{\partial}{\partial x}\left[x^T A^T A x - x^T A^T b - b^T A x + v^{*T} C x\right]$$

Lagrangian : $L(x, v) = \underbrace{(Ax - b)^T (Ax - b)}_{f(x)} + v^T \underbrace{(Cx - d)}$

KKT conditions:

$\nabla_x(L(x, v^*)) = 2A^T Ax - 2A^T b + C^T v^*$

$\nabla_x L(x, v^*) = 0$ at $x = x^*$

$2A^T Ax^* - 2A^T b + C^T v^* = 0$ — ①

Solution :

$$\begin{bmatrix} 2A^T A & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} 2A^T b \\ d \end{bmatrix}$$

KKT Matrix

Solve for x^* & v^* together as $\begin{bmatrix} x^* \\ v^* \end{bmatrix} = \begin{bmatrix} 2A^T A & C^T \\ C & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2A^T b \\ d \end{bmatrix}$

KKT Matrix is square and is invertible, if solution exists

Constrained Least Squares: Advertising Purchase Example

- m demographic groups of audiences
- n number channels to advertise
- $m \times n$ Matrix R represents the 'Available Data' on Ad views per dollar spent
- v^{des} is the desired viewership from each channel
- $n -$ vector s is the dollars invested in each channel for advertisement
- $m -$ vector $Rs = v$ gives the total viewership from each demographic group

$f_0(x)$

s.t. $a_i^T x = b$

$n = 3$ channels

$m = 10$ demographic groups

units : 1000 views per dollar

$$\begin{array}{cc} & \begin{array}{ccc} (s_1) & (s_2) & (s_3) \end{array} \\ 10 & R = \begin{bmatrix} 0.97 & 1.86 & 0.41 \\ 1.23 & 2.18 & 0.53 \\ 0.80 & 1.24 & 0.62 \\ 1.29 & 0.98 & 0.51 \\ 1.10 & 1.23 & 0.69 \\ 0.67 & 0.34 & 0.54 \\ 0.87 & 0.26 & 0.62 \\ 1.10 & 0.16 & 0.48 \\ 1.92 & 0.22 & 0.71 \\ 1.29 & 0.12 & 0.62 \end{bmatrix} \end{array}$$

v_{des}

$$\begin{bmatrix} 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \\ 10^3 \end{bmatrix}$$

$v^{des} = (10^3)\mathbf{1}$

Unconstrained Objective:

Find $\hat{s}$ that minimizes $\|Rs - v^{des}\|^2$

$$\hat{s} = \begin{bmatrix} 62 \\ 100 \\ 1443 \end{bmatrix}$$

Total Dollars to be spent is 1605

Constrained Objective:

$$\begin{array}{ll} \text{minimize} & \|Rs - v^{des}\|^2 \\ \text{subject to} & \mathbf{1}^T s = B. \end{array}$$

$c \qquad d$

B is the total budget available

Actual budget available is 20% less i.e. B = 1284

Solving constrained least squares, we get

$$\hat{s} = \begin{bmatrix} 315 \\ 110 \\ 859 \end{bmatrix}$$