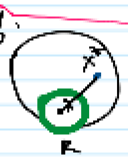
Special HW 5&6

HW6: L7 - page 2

Global Optimum: $x^* \in \mathbb{R}^d$ is a global optimum
if x^*'s feasible &
$g(x^*) = \inf\{g(z) \mid \forall \text{ feasible } z\}$

$$\boxed{\begin{array}{l}\text{For a convex function optimization over a} \\ \text{convex feasible set.} \\ \qquad \text{local optimum} \Rightarrow \text{global optimum}\end{array}}$$

Proof:



x is a local optimum
Suppose x is not a global optimum.
let x^* be the global optimum.
x is local optimum: $\exists$ a radius R s.t.

x is local optimum: $\exists$ a radius R s.t.
$x = \inf\{y \mid \|y - x\|_2 \le R,\ y \text{ is feasible}\}$

Consider the line segment: x — x^*.
Consider a point. $y = \theta x^* + (1-\theta)x$

x —— x^* $\quad \|x^* - x\|_2$

When $\theta = \dfrac{R}{2\|x^* - x\|_2}$, y satisfies $\|y - x\|_2 \le R$

HW: Show that is true

By convexity of $g(\,)$.
$g(y) \le \theta g(x^*) + (1-\theta) g(x)$.
$= g(x) - \theta(g(x) - g(x^*))$

$g(x^*) < g(x) \Rightarrow \underline{g(x) - g(x^*) \ge 0}$
$\theta \ge 0$.
$\Rightarrow Z = \theta(g(x) - g(x^*)) \ge 0$.

$\Rightarrow g(y) = g(x) - Z < g(x)$
Contradicts the assumption that x is a local optimum & not a global optimum.

Statement:

For the optimization of a convex function over a convex feasible set, local optimum = global optimum. x is a local optimum. Suppose x is not a global optimum. Let x* be the global optimum. x is a local optimum.... x = inf {y | ||y-x||| <= R, y is feasible}. Consider the line segment x --> x*. Consider a point y = tetha x* + (1-theta)x. when tetha = R ('radius') / (2 |||x* - x||), y satisfies ||y -x|| <= R... prove that: y satisfies ||y -x|| <= R → IS TRUE.

Solution:

We have to show that $||y - x|| \leq R$, where $y = \theta x^* + (1 - \theta)x$, $\theta = R/(2|||x^* - x||)$, and x* is the global optimum.

First, note that $|||x^* - x|| > 0$ since x is a local but not global optimum, and x* is the global optimum.

Next, we have:

$||y - x|| = ||\theta x^* + (1 - \theta)x - x|| = ||\theta(x^* - x)|| = \theta\ ||x^* - x||$.

Substituting $\theta = R/(2|||x^* - x||)$, we obtain:

$||y - x|| = (R/(2||x^* - x||))\ ||x^* - x|| = R/2 < R$

Therefore, y satisfies $||y - x|| <= R$, which completes the proof.

HW7: L8 - page 1

CS538-L8

Monday, February 6, 2023 5:04 PM

Solving Convex Programs:

A

$$\boxed{\begin{array}{ll} \min & g(x) \\ \text{s.t.} & f_i(x) \le 0 \quad i = 1, \dots, m' \\ & h_j(x) = 0 \quad j = m'+1, \dots, m \\ & x \in \mathbb{R}^n \end{array}}$$

$$\min \; x_1^2 + 2x_2^2 + x_3^2$$

s.t.

λ_1: $x_1 + 3x_2 = 5$

λ_2: $3x_1 + x_3 = 7$

$$(y_1 y_2 y_3)\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$2y_1^2 + 4y_2^2 + 2y_3^2$$

always ≥ 0.

Convexity

Lagrangian Method:

$$L(x, \lambda_1, \lambda_2) = x_1^2 + 2x_2^2 + x_3^2 + \lambda_1(x_1 + 3x_2 - 5) + \lambda_2(3x_1 + x_3 - 7)$$

For equality constraints: m constraints.

$$\boxed{\min \; g(x) \quad \text{s.t.} \quad h_j(x) = 0, \; j = 1 \dots m}$$

$$L(x, \lambda) = g(x) + \sum_j \lambda_j h_j(x)$$

$$x \in \mathbb{R}^n, \; \lambda \in \mathbb{R}^m$$

$$\min \; x_1^2 + 2x_2^2 + x_3^2 \quad \text{s.t.} \quad \underbrace{x_1 + 3x_2 - 5 = 0}_{\lambda_1}; \; \underbrace{3x_1 + x_3 - 7 = 0}_{\lambda_2}$$

$$L(x, \lambda) = x_1^2 + 2x_2^2 + x_3^2 + \lambda_1(x_1 + 3x_2 - 5) + \lambda_2(3x_1 + x_3 - 7)$$

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

$$\left.\begin{array}{l} \frac{\partial L}{\partial \lambda_1}: \; x_1 + 3x_2 - 5 = 0 \\ \frac{\partial L}{\partial \lambda_2}: \; 3x_1 + x_3 - 7 = 0 \\ \frac{\partial L}{\partial x_1}: \; 2x_1 + \lambda_1 + 3\lambda_2 = 0 \\ \partial L/\partial x_2: \; 4x_2 + 3\lambda_1 + 0 = 0 \\ \partial L/\partial x_3: \; 2x_3 + 0 + \lambda_2 = 0 \end{array}\right\}$$

5 equations,
5 unknowns

HW: Solve for minimum point.

<u>Statement:</u>

We have the following system of equations:

dL/dλ1 = 0 = x1 + 3*x2 - 5
dL/dλ2 = 0 = 3*x1 + x3 - 7
dL/dx1 = 0 = 2*x1 + λ1 + 3*λ2

dL/dx2 = 0 = 4*x2 + 3*λ1
dL/dx3 = 0 = 2*x3 + λ2

Solution:

Using the fourth equation, λ1 can be solved in terms of x2

dL/dx2 = 0 = 4*x2 + 3*λ1 => λ1 = -(4/3)*x2

With the fifth equation, λ2 can be obtained in terms of x3:

dL/dx3 = 0 = 2*x3 + λ2 => λ2 = -2*x3

Substituting these expressions into the third equation we get rid of the lambda parameters:

dL/dx1 = 0 = 2*x1 - (4/3)*x2 - 6*x3

Now obtaining x1 as a function of x2, x3

2*x1 = (4/3)*x2 + 6*x3 => x1 = (2/3)*x2 + 3*x3

Substituting these expressions for x1, λ1 , λ2 into the first equation, we obtain:

dL/dλ1 = 0 = (2/3)*x2 + 9*x3 - 5… if we keep operating between equations, we just solve the feasible system of 5 equations and 5 variables, whose results are as follows:

The minimum point of the function is:

x1 = ***2.16***
x2 = 87 / 92 = ***0.94***
x3 = 47 / 92 = ***0.51***
λ1 = -(29 / 23) = ***- 1.26***
λ2 = -(47/46) = ***- 1.02***

Flow Problem

$$\boxed{\begin{aligned} \max\ & \sum_{p \in \mathcal{P}} f_p \\ \text{s.t.}\ & Hf \le C \\ & f \ge 0 \end{aligned}}$$

$f^T = (f_{P_1}, f_{P_2}, f_{P_3}, \cdots)$ $\quad P_i \in \mathcal{P}$

C — capacity of edges vector $\begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$

Solution in example: $f_{P_1} = 1, f_{P_2} = 0$
$f_{P_3} = 1$

$\vec{e} = (1, \cdots)$

$$\boxed{\begin{aligned} \max\ & e^T f \\ \text{s.t.}\ & C - Hf \ge 0 \quad \lambda \\ & f \ge 0 \quad \mu \end{aligned}} \xleftrightarrow{\text{Dual}} \boxed{??}$$

dual variable λ, μ $(\lambda \ge 0, \mu \ge 0)$

$$L(f, \lambda, \mu) = e^T f + \lambda^T (C - Hf) + \mu^T f$$

HW use summation notation

$$\phi(\lambda, \mu) = \sup_f \left[e^T f + \lambda^T (C - Hf) + \mu^T f \right]$$

$$= \lambda^T C + \sup_f \left[\underbrace{(e^T - \lambda^T H + \mu^T)}\, f \right]$$

$\lambda \ge 0, \mu \ge 0$.

This implies a condition

$$e^T - \lambda^T H + \mu^T = 0.$$

Dual Program:

$$\boxed{\begin{aligned} \min\ & \lambda^T C \\ \text{s.t.}\ & e^T - \lambda^T H + \mu^T = 0, \lambda \ge 0, \mu \ge 0 \end{aligned}}$$

$$\equiv \min \lambda^T c$$
$$H^T \lambda = e + \mu, \ \mu \ge 0, \lambda \ge 0$$

$$\equiv \boxed{\begin{aligned} \min\ & \underline{C^T \lambda} \\ & H^T \lambda \ge e, \lambda \ge 0 \end{aligned}}$$

$$\begin{array}{cc} & \text{edges} \\ H^T = \text{path} & \left[\quad \right] \end{array} \begin{pmatrix} \lambda_e \end{pmatrix} \ge \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$|||$$

$$\boxed{\begin{aligned} \min\ & \sum_e c_e \lambda_e \\ \text{s.t.}\ & \sum_{e \in p} \lambda_e \ge 1, \forall p \end{aligned}} \equiv \underline{\text{Min-Cut}}$$

HW justify

$$\boxed{\lambda_e = \begin{cases} 1 & e \text{ in solution} \\ 0 & e \text{ not in solution} \end{cases}}$$

A feasible λ which is $\{0, 1\}$ identifies a cut.

By strong duality LP solution = Dual LP solution.

$\Rightarrow$ Max-Flow = Min-Cut

1) ***Statement:***

Use the summation notation for representing the next equation:

$$\boldsymbol{L(f, \lambda, v) = \varphi T\ f_i + \lambda^T (c - H\ f) + \mu^T\ f}$$

1) ***Solution:***

The flow problem with the given function can be expressed using the summation notation as follows:

$$\boldsymbol{L(f, \lambda, v) = \Sigma_i\ \varphi_i\ f_i + \Sigma_j\ \lambda_j (c_j - \Sigma_i\ H_{ji}\ f_i) + \Sigma_i\ v_i\ f_i}$$

Where:

Σ_i denotes the sum over all indices i.
Σ_j is the sum over all indices j.
φ is a vector of coefficients of the flux variable f.
λ is a vector of Lagrange multipliers associated with the constraint c - Hf = 0.
v is a vector of Lagrange multipliers associated with the nonnegativity constraint of f.
c is a vector of constants.
H is a matrix of coefficients.
f is the flux variable.
Note that the dot product between two vectors can be expressed using summation notation as $\Sigma_i\ x_i y_i$. In this case, $\Sigma_i\ \varphi_i\ f_i$ represents the scalar product between the vectors φ and f. Similarly, $\Sigma_i\ H_{ji}\ f_i$ represents the scalar product between the j-th row of H and f.

2) ***Statement:***

Justify that the min cut problem is the dual of the flow problem

2) ***Solution:***

The duality of the flow problem:

The flow problem can be expressed mathematically as follows:

Maximize Z = ∑ fij, i, j ∈ N.

Subject to:

fij ≤ cij for all i, j ∈ N.
∑ fij - ∑ fji = 0 for all i ∈ N
0 ≤ fij ≤ uij for all i, j ∈ N

where fij represents the flow from node i to node j, cij represents the capacity of the edge between nodes i and j, and uij represents the upper bound of the flow between nodes i and j. The objective function Z represents the total flow over the network of N nodes.

On the other hand, the min cut problem can be expressed mathematically as follows:

Minimize $Z = \sum c_{ij}$, $i, j \in N$.

Subject to:

$S_i \subseteq N$, $T_i \subseteq N$, $S_i \cap T_i = \varnothing$, and $|S_i| > 0$, $|T_i| > 0$.
$\sum f_{ij} - \sum f_{ji} = 0$ for all $i \in N$
$f_{ij} \leq c_{ij}$ for all $i, j \in N$
$f_{ij} \geq 0$ for all $i, j \in N$

where S_i and T_i represent two disjoint sets of nodes partitioning the network, and c_{ij} represents the cost of cutting the edge between nodes i and j. The objective function Z represents the total cost associated with splitting the network into two parts.

The min cut problem is the dual of the flow problem due to the fact that the two problems are related through duality: any feasible flow in the flow problem can be represented as a cut in the min cut problem, and any network cut can be represented as constraints in the flow problem.

The flow problem is a linear programming problem in which flow variables must be limited so that the total flow over a network of nodes can be maximized. In contrast, the min cut problem seeks to minimize the total cost of dividing the network into two parts within certain cut variables constraints.

The duality of the two problems makes it possible to resolve one by resolving the other's dual, or the other way. This tool is really helpful because sometimes, the dual problem is easier than the main one. Despite the fact that they are connected by duality, it is important to note that the approaches to solving these problems may be different.