Spring 2023

SWCON253: Machine Learning

Lecture 12
Constrained Optimization

Jinwoo Choi
Assistant Professor
CSE, Kyung Hee University

Vision & **Learning** Lab
Kyung Hee University

Contents

References

- *Chapter 7, Mathematics for Machine Learning* by Deisenroth, Faisal, and Ong (https://mml-book.com)
- *Intro to Deep Learning & Generative Models* by Sebastian Raschka (http://pages.stat.wisc.edu/~sraschka/teaching/stat453-ss2020/)
- *패턴 인식* by 오일석, *기계 학습* by 오일석

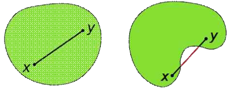
(Cf.) Convex Set & Convex Function

Convex Set

Definition 7.2. A set $\mathcal{C}$ is a *convex set* if for any $x, y \in \mathcal{C}$ and for any scalar θ with $0 \leqslant \theta \leqslant 1$, we have

$$\theta x + (1-\theta)y \in \mathcal{C}\,. \tag{7.29}$$

Linear segment btw two points in C lies in C.



↳ 오목한 부분이 없어야 함

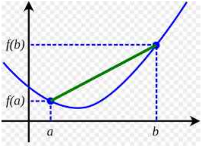
Convex Function

Definition 7.3. Let function $f : \mathbb{R}^D \to \mathbb{R}$ be a function whose domain is a convex set. The function f is a *convex function* if for all $\boldsymbol{x}, \boldsymbol{y}$ in the domain of f, and for any scalar θ with $0 \leqslant \theta \leqslant 1$, we have

$$\boxed{f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leqslant \theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})}\,. \tag{7.30}$$

Remark. A *concave function* is the negative of a convex function. ◊

A straight line between any two points of the function lie above the function.



NOTE:

All the derivations in this lecture assume ***convexity*** of the variables & functions.

볼록성 보장

1. *Constrained* Optimization Problems

- **With *No* Constraints**

$\boxed{\text{minimize } f(x)}$

- *Unconstrained* Optimization
- Solution:

$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \mathbf{0}$

- **With *Equality* Constraints**

등식 제한 조건

$$\boxed{\begin{array}{l}\text{minimize } f(x) \\ \text{subject to } g(x) = 0\end{array}}$$

모든 연립방정식 g를 만족하며 f를 최소화 하는 x

- Example

$$f(x_1, x_2) = x_1^2 + x_2^2 - 1$$
$$g(x_1, x_2) = x_1 + x_2 - 1$$

- **With *Inequality* Constraints**

$$\boxed{\begin{array}{l}\text{minimize } f(x) \\ \text{subject to } g(x) \leq 0\end{array}}$$

- Example

maximize $x \cdot y$
subject to $2x + 2y \leq 1$

⬇

minimize$(-x \cdot y)$
subject to $2x + 2y - 1 \leq 0$

2. Lagrange Multiplier for *Equality* Constraints

◆ An Illustrative Example

minimize $f(\underline{x})$
subject to $g(\underline{x}) = 0$

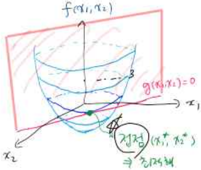
$$f(x_1, x_2) = x_1^2 + x_2^2 - 1$$ (제약 조건 없으면 (0,0))

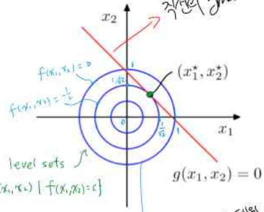
$$g(x_1, x_2) = x_1 + x_2 - 1$$

ex) $x+y-1=0$

$\nabla = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

직선의 gradient는 직선의 수직방향



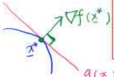


$f(x_1, x_2) = C$ 인 동심원

C가 작은 점을 찾자 -

At the optimal point $\boldsymbol{x}^*$, the gradient of f is orthogonal to the surface $\{g(\boldsymbol{x}) = 0\}$.

Consequently, $\nabla g(\boldsymbol{x})$ & $\nabla f(\boldsymbol{x})$ are parallel at $\boldsymbol{x}^*$.



∇g & ∇f

"parallel"

$\nabla f + \lambda \nabla g = 0$

새롭게 $L = f + \lambda g$ 로 정의.

$\nabla L = 0$ 찾자

2. Lagrange Multiplier for *Equality* Constraints (cont'd)

Equality Constrained Optimization

minimize $f(\underset{\sim}{x})$
subject to $g(\underset{\sim}{x}) = 0$

- At the optimal point $\boldsymbol{x}^*$, $\nabla_{\boldsymbol{x}} g(\boldsymbol{x})$ and $\nabla_{\boldsymbol{x}} f(\boldsymbol{x})$ are *parallel*.

  ★ Hence, there exists some $\lambda \in \boldsymbol{R}$ such that $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \lambda \nabla_{\boldsymbol{x}} g(\boldsymbol{x}) = \mathbf{0}$.

- We define the *Lagrangian* function $L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) + \lambda g(\boldsymbol{x})$,

  where $\lambda \in \boldsymbol{R}$ is called *Lagrangian multiplier*.

- Now observe that:

  ★ $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) + \lambda \nabla_{\boldsymbol{x}} g(\boldsymbol{x}) = \mathbf{0} \quad \Leftrightarrow \quad \nabla_{\boldsymbol{x}} L(\boldsymbol{x}, \lambda) = \mathbf{0}$

  ★ $g(\boldsymbol{x}) = 0 \quad \Leftrightarrow \quad \nabla_{\lambda} L(\boldsymbol{x}, \lambda) = 0$

Example

$$f(x_1, x_2) = x_1^2 + x_2^2 - 1$$
$$g(x_1, x_2) = x_1 + x_2 - 1$$

$$L(\boldsymbol{x}, \lambda) = x_1^2 + x_2^2 - 1 + \lambda(x_1 + x_2 - 1)$$

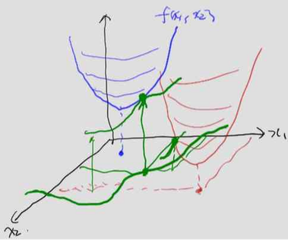
$$\nabla_{x_1} L(\boldsymbol{x}, \lambda) = 2x_1 + \lambda = 0$$
$$\nabla_{x_2} L(\boldsymbol{x}, \lambda) = 2x_2 + \lambda = 0$$
$$\nabla_{\lambda} L(\boldsymbol{x}, \lambda) = x_1 + x_2 - 1 = 0$$

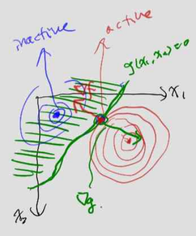
$$\therefore (x_1^*, x_2^*) = (0.5, 0.5)$$

3. Lagrange Multiplier for *Inequality* Constraints (cont'd)

$$\min.\ f(x_1, x_2)$$
$$s.t.\ g(x_1, x_2) \le 0$$

- ∇□(□ □□의 방향: 제약조건이 g(x)< 0 이므로 g의 값은 경계(g(x)= 0)에서 최대
  → 경계 안쪽(g(x)< 0)으로 갈수록 g값이 경계보다 줄어듦 → g의 증가방향은 경계 바깥 방향
- 제약조건이 active: 제약조건을 만족하는 f의 최소값(최적해)이 경계(g(x)= 0)에 존재한다는 뜻
  → f의 원래 최소값이 경계 바깥쪽(g(x)≥ 0)에 존재 → f의 증가방향은 경계 안쪽 방향
- 제약조건이 inactive: 제약조건을 만족하는 f의 최소값(최적해)이 경계 안쪽(g(x)< 0)에 존재한다는 뜻
  → f의 원래 최소값이 경계 안쪽 (참고: f의 증가 방향은 경계 바깥 방향)





At the optimal point □□,

1) If active:

- ∇□(□ □□ points *outwards*
- ∇□(□ □□ points *inwards*

2) If inactive:

- ∇□(□ □□ = □
  (unconstrained optimization)

3. Lagrange Multiplier for *Inequality* Constraints

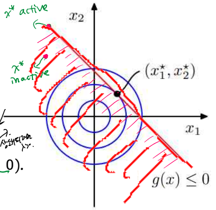
- ***Inequality*** **Constrained Optimization**

minimize $f(x)$
subject to $g(x) \le 0$

$$f(x_1, x_2) = x_1^2 + x_2^2 - 1$$
$$g(x_1, x_2) = x_1 + x_2 - 1$$

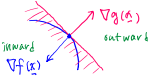
- Case 1: Constraint is "***active***" (i.e., x^* is on $g(x) = 0$)
  - Again, $\nabla_x f(x)$ and $\nabla_x g(x)$ are *parallel* at the optimal point x^*
  - Furthermore, $\nabla_x g(x)$ points *outwards* and $\nabla_x f(x)$ points *inwards*
  - So we have: $\nabla_x f(x) + \lambda \nabla_x g(x) = 0$ for some $\lambda > 0$. 왜?

f(x)의 grad는 밖에 있고, 기울기가
따라서 λ에는 상쇄되려면 λ>0.



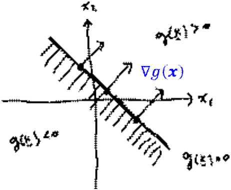
- Case 2: Constraint is "***inactive***" (i.e., x^* is ***not*** on $g(x) = 0$)
  - In this case, we can treat the problem as ***unconstrained*** optimization ($\lambda = 0$).
  - Thus, we have $\nabla_x f(x) = 0$ at the optimal point x^*

⇒
- Active $(g(x) = 0)$: $\nabla_x L(x, \lambda) = 0$ for some $\lambda > 0$
- Inactive $(g(x) < 0)$: $\nabla_x L(x, \lambda) = 0$ for $\lambda = 0$.

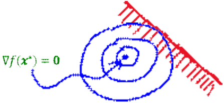


3. Lagrange Multiplier for *Inequality* Constraints (cont'd)

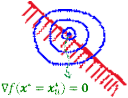
- **[Note 1] Gradient of g**
  - Inequality constraint: $g(\boldsymbol{x}) \leq 0$
  - $\nabla g(\boldsymbol{x})$ is *outward* the boundary
    - ★ since at the boundary $g(\boldsymbol{x})$ increases on the outward direction (from negative to positive).

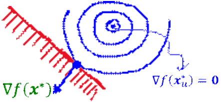


- **[Note 2] Gradient of *f***
  - **Inactive case** (i.e., solution is *inside* the boundary)
    - ★ The unconstrained solution must reside *inside* the boundary → $\nabla f(\boldsymbol{x}^*) = \mathbf{0}\ (\boldsymbol{x}^* = \boldsymbol{x}^*_{unc})$



  - **Active case** (i.e., solution is on the boundary)
    - ★ The unconstrained solution must reside either *on* or *outside* the boundary → $\nabla f(\boldsymbol{x}^*)$ is either **zero** or ***inward*** the boundary.





3. Lagrange Multiplier for *Inequality* Constraints (cont'd)

◆ *Inequality* Constrained Optimization (cont'd)

- We can summarize both cases using the Lagrangian again.

$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) + \lambda g(\boldsymbol{x})$

$\boxed{\nabla_{\boldsymbol{x}} L(\boldsymbol{x}, \lambda) = \boldsymbol{0}}$ (stationarity or saddle point condition)

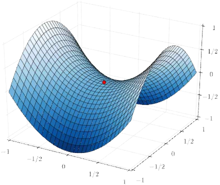
$g(\boldsymbol{x}) \le 0$ (primal feasibility)

$\lambda \ge 0$ (dual feasibility)

$\lambda\, g(\boldsymbol{x}) = 0$ (complementary slackness)

- $\boxed{g(\boldsymbol{x}) = 0}$, $\lambda > 0$ (active) $\Rightarrow$ $\boxed{\nabla_{\lambda} L(\boldsymbol{x}, \lambda) = 0}$
- $g(\boldsymbol{x}) < 0$, $\lambda = 0$ (inactive)

A ***saddle point*** on the graph of $z=x^2-y^2$ (hyperbolic paraboloid)



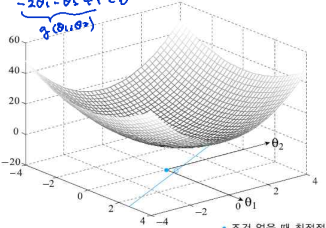
https://en.wikipedia.org/wiki/Saddle_point#/media/File:Saddle_point.svg

- The conditions in the yellow box is called the ***Karush-Kuhn-Tucker (KKT) condition***
  - ★ It is sometimes also called as *first-order necessary condition*.

<> Examples – *Equality Constraint*

- 등식 조건부 최적화 문제: $J(\boldsymbol{\theta}) = \theta_1^2 + 2\theta_2^2$
  - 조건이 없을 때 최소 점은 $(0,0)^T$
  - '$2\theta_1 + \theta_2 = 1$' 이라는 조건 하에 최소점을 구하면?

$\underbrace{-2\theta_1 - \theta_2 + 1}_{g(\theta_1, \theta_2)} = 0$



- 조건 없을 때 최적점$(0, 0)^T$
- 조건 있을 때 최적점$(4/9, 1/9)^T$

그림 11.11 등식 조건부 최적화 문제의 예

$g(\theta_1, \theta_2)$

$\mathbin{/\!/}$

라그랑제 함수: $L(\boldsymbol{\theta}, \lambda_1) = (\theta_1^2 + 2\theta_2^2) - \lambda_1 (2\theta_1 + \theta_2 - 1)$

$\boldsymbol{\theta}$로 미분한 식을 $\mathbf{0}$으로 둠: $\begin{cases} \dfrac{\partial L}{\partial \theta_1} = 2\theta_1 - 2\lambda_1 = 0 \\ \dfrac{\partial L}{\partial \theta_2} = 4\theta_2 - \lambda_1 = 0 \end{cases}$

λ로 미분한 식을 0으로 둠: $\dfrac{\partial L}{\partial \lambda_1} = -(2\theta_1 + \theta_2 - 1) = 0$

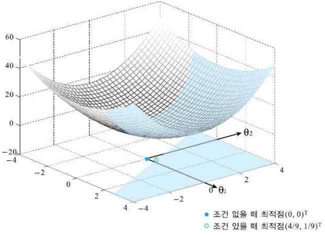
이 식을 풀면, $\hat{\boldsymbol{\theta}} = \left(\dfrac{4}{9}, \dfrac{1}{9}\right)^T$, $\lambda = \dfrac{4}{9} > 0$

<> Examples – *Inequality Constraint*

◆ 부등식 조건부 최적화 문제:

$J(\boldsymbol{\theta}) = \theta_1^2 + 2\theta_2^2$

● '$2\theta_1 + \theta_2 \geq 1$'이라는 조건 하에 최소점을 구하라.



(a) $f_1(\boldsymbol{\theta}) = 2\theta_1 + \theta_2 - 1 \geq 0$

라그랑제 함수: $L(\boldsymbol{\theta}, \lambda_1) = (\theta_1^2 + 2\theta_2^2) - \lambda_1 \overbrace{(2\theta_1 + \theta_2 - 1)}^{g(\theta)}$

KKT 조건:
$$\begin{cases} \dfrac{\partial L}{\partial \theta_1} = 2\theta_1 - 2\lambda_1 = 0 \\ \dfrac{\partial L}{\partial \theta_2} = 4\theta_2 - \lambda_1 = 0 \\ \lambda_1 \geq 0 \\ \lambda_1(2\theta_1 + \theta_2 - 1) = 0 \end{cases}$$

이제 KKT 조건을 풀어 이를 만족하는 해 $\boldsymbol{\theta}$ 를 구하면 된다. 이 예는 라그랑제 승수를 하나만 가지므로 쉽게 풀 수 있다. 마지막 식에서 $\lambda_1 = 0$이거나 $2\theta_1 + \theta_2 - 1 = 0$이어야 한다. 먼저 $\lambda_1 = 0$이라고 가정해 보자. 그럼 $\theta_1 = \theta_2 = 0$이 되어 주어진 조건 $f_1(\boldsymbol{\theta}) = 2\theta_1 + \theta_2 - 1 \geq 0$을 만족하지 못한다. 두 번째 경우를 가지고 풀어 보면 $\hat{\boldsymbol{\theta}} = (\frac{4}{9}, \frac{1}{9})^{\mathrm{T}}$을 얻는다.

① active : $\boxed{\nabla_\lambda L = 0} \Rightarrow -2\theta_1 - \theta_2 + 1 = 0$

solution : $\underline{\theta} = (\frac{4}{9}, \frac{1}{9})^T$, $\lambda = \frac{4}{9} > 0$, (O)

② inactive : $\boxed{\lambda = 0} \Rightarrow \underline{\theta} = \underline{0} \Rightarrow g(\underline{\theta}) = 1 > 0$ (X)

4. Lagrange Multiplier for *Multiple* Constraints

- **Multiple Constraints** (*including both equality & inequality constraints*)

$$
\boxed{\begin{aligned} &\text{minimize } f_0(\underline{x}) \\ &\text{subject to } f_i(\underline{x}) \le 0 \quad (i = 1, \dots, m) \\ &\qquad\qquad\quad h_j(\underline{x}) = 0 \quad (j = 1, \dots, k) \end{aligned}}
$$

제한 조건 여러개

$$
L(\underline{x}, \underline{\lambda}, \underline{\nu}) = f_0(\underline{x}) + \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) + \sum_{j=1}^{k} \nu_j h_j(\underline{x}) = f_0(\underline{x}) + \underline{\lambda}^T \underline{f}(\underline{x}) + \underline{\nu}^T \underline{h}(\underline{x})
$$

★ λ_i : Lagrange multiplier associated with $f_i(x_i) \le 0.$

★ ν_i : Lagrange multiplier associated with $h_i(x_i) = 0.$

- Note: Let $\boldsymbol{x}_0$ be a feasible point of the primal problem
(that is a point that satisfies all the constraints), then we have:

$$
\sum_{i=1}^{m} \underbrace{\lambda_i}_{\ge 0} \underbrace{f_i(x_0)}_{\le 0} + \sum_{i=1}^{k} \nu_j \underbrace{h_j(x_0)}_{=0} \le 0 \quad \Rightarrow \quad L(x_0, \lambda, \nu) = f_0(x_0) + \sum_{i=1}^{m} \lambda_i f_i(x_0) + \sum_{j=1}^{k} \nu_j h_j(x_0) \le f_0(x_0)
$$

5. Lagrangian Duality

- **Duality in Optimization**
  - The idea of converting an optimization problem in one set of variables $\boldsymbol{x}$ (called the *primal variables*) into another set of variables $\boldsymbol{\lambda}$ (called the *dual variables*).

- **Primal Problem**

$$\boxed{\begin{aligned} &\min_{\boldsymbol{x}} \quad f(\boldsymbol{x}) \\ &\text{subject to} \quad g_i(\boldsymbol{x}) \leqslant 0 \quad \text{for all} \quad i = 1, \ldots, m \end{aligned}}$$

- **Lagrangian Dual Problem**

$$\boxed{\begin{aligned} &\max_{\boldsymbol{\lambda} \in \mathbb{R}^m} \quad \mathfrak{D}(\boldsymbol{\lambda}) = \boxed{\min_{\boldsymbol{x} \in \mathbb{R}^d} \mathfrak{L}(\boldsymbol{x}, \boldsymbol{\lambda})} \\ &\text{subject to} \quad \boldsymbol{\lambda} \geqslant 0\,, \end{aligned}}$$

$\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$

$\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow x^* = \ldots$

$\therefore D(\lambda) = \mathcal{L}(x^*, \lambda)$

5. Lagrangian Duality – *Linear Programming*

Consider the special case when all the preceding functions are linear, i.e.,

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^d} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\,, \end{aligned} \tag{7.39}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times d}$ and $\boldsymbol{b} \in \mathbb{R}^m$.

The Lagrangian is given by

$$\mathfrak{L}(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{\lambda}^\top(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = (\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda})^\top \boldsymbol{x} - \boldsymbol{\lambda}^\top \boldsymbol{b}\,.$$

Taking the derivative of $\mathfrak{L}(\boldsymbol{x},\boldsymbol{\lambda})$ with respect to $\boldsymbol{x}$ and setting it to zero $\left({}_{\boldsymbol{x}}\mathfrak{L}(\boldsymbol{x},\boldsymbol{\lambda}) = \mathbf{0} \right)$

$$\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda} = \mathbf{0}.$$

$= \mathfrak{L}(x, \lambda)$

Therefore, the dual Lagrangian is $\mathfrak{D}(\boldsymbol{\lambda}) = -\boldsymbol{\lambda}^\top\boldsymbol{b}$. $\mathfrak{D}(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}\in\mathbb{R}^d} \mathfrak{L}(\boldsymbol{x},\boldsymbol{\lambda})$.

dual optimization problem

$$\begin{aligned} \max_{\boldsymbol{\lambda}\in\mathbb{R}^m} \quad & -\boldsymbol{b}^\top\boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda} = \mathbf{0} \\ & \boldsymbol{\lambda} \geqslant \mathbf{0}. \end{aligned}$$

We have the choice of solving the primal or the dual program depending on whether *m* or *d* is larger.

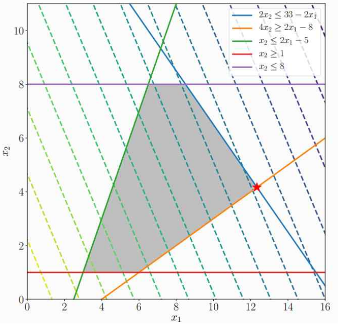
5. Lagrangian Duality – *Linear Programming (cont'd)*

Example 7.5 (Linear Program)
Consider the linear program

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \quad -\begin{bmatrix}5\\3\end{bmatrix}^\top \begin{bmatrix}x_1\\x_2\end{bmatrix}$$

$$\text{subject to} \quad \begin{bmatrix}2 & 2\\ 2 & -4\\ -2 & 1\\ 0 & -1\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} \leqslant \begin{bmatrix}33\\8\\5\\-1\\8\end{bmatrix} \tag{7.44}$$

with two variables. This program is also shown in Figure 7.9. The objective function is linear, resulting in linear contour lines. The constraint set in standard form is translated into the legend. The optimal value must lie in the shaded (feasible) region, and is indicated by the star.



해보자

5. Lagrangian Duality – *Quadratic Programming*

Consider the case of a convex quadratic objctive function, where the constraints are affine, i.e.,

$$\min_{\boldsymbol{x}\in\mathbb{R}^d} \quad \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^\top \boldsymbol{x} \tag{7.45}$$
$$\text{subject to} \quad \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\,,$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times d}$, $\boldsymbol{b} \in \mathbb{R}^m$, and $\boldsymbol{c} \in \mathbb{R}^d$.

The square symmetric matrix $\boldsymbol{Q} \in \mathbb{R}^{d\times d}$ is positive definite, and therefore the objective function is convex.

The Lagrangian is given by $\mathfrak{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{\lambda}^\top(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + (\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda})^\top \boldsymbol{x} - \boldsymbol{\lambda}^\top \boldsymbol{b}\,,$

Taking the derivative of $\mathfrak{L}(\boldsymbol{x}, \boldsymbol{\lambda})$ with respect to $\boldsymbol{x}$ and setting it to zero gives $\boldsymbol{Q}\boldsymbol{x} + (\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda}) = \boldsymbol{0}\,.$

Assuming that $\boldsymbol{Q}$ is invertible, we get $\boldsymbol{x} = -\boldsymbol{Q}^{-1}(\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda})\,.$

Therefore, the dual optimization problem is given by

$$\max_{\boldsymbol{\lambda}\in\mathbb{R}^m} \quad -\frac{1}{2}(\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda})^\top \boldsymbol{Q}^{-1}(\boldsymbol{c} + \boldsymbol{A}^\top\boldsymbol{\lambda}) - \boldsymbol{\lambda}^\top \boldsymbol{b}$$
$$\text{subject to} \quad \boldsymbol{\lambda} \geqslant \boldsymbol{0}\,.$$

We have the choice of solving the primal or the dual program depending on whether *m* or *d* is larger.

5. Lagrangian Duality – *Quadratic Programming (cont'd)*

Example: Converting into the standard QP form and then solving the dual problem

- **Primal Problem**

$$\boxed{\begin{aligned} \min_{\boldsymbol{x}} \quad & J(\boldsymbol{x}) = x_1^2 + 2x_2^2 \\ \text{s.t.} \quad & g(\boldsymbol{x}) = -2x_1 - x_2 + 1 \leq 0 \end{aligned}}$$

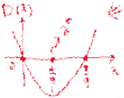
$$\Rightarrow \quad \boxed{\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^d} \quad & \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}, \end{aligned}} \quad \rightarrow J(x)$$

$\longrightarrow \; Ax - b \leq 0 \;\Rightarrow\; g(x) = Ax - b$

Let $Q = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, $c = 0$, $A = [-2 \; -1]$, $b = -1$

Then we can represent the original problem in the standard QP form.

- **Lagrangian Dual Problem**



$$\Leftarrow \quad \begin{aligned} \max_{\lambda} \quad & -\frac{9}{8}\lambda^2 + \lambda \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned} \quad \Leftarrow \quad \boxed{\begin{aligned} \max_{\boldsymbol{\lambda}\in\mathbb{R}^m} \quad & -\frac{1}{2}(\boldsymbol{c}+\boldsymbol{A}^\top\boldsymbol{\lambda})^\top \boldsymbol{Q}^{-1}(\boldsymbol{c}+\boldsymbol{A}^\top\boldsymbol{\lambda}) - \boldsymbol{\lambda}^\top \boldsymbol{b} \\ \text{subject to} \quad & \boldsymbol{\lambda} \geqslant \boldsymbol{0}. \end{aligned}}$$

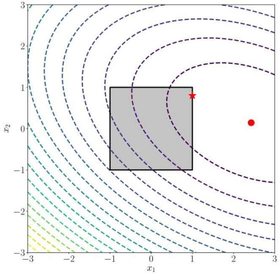
$\Rightarrow$ Sol. $\lambda^* = \frac{4}{9}$ $\Rightarrow$ $\boldsymbol{x} = -\boldsymbol{Q}^{-1}(\boldsymbol{c}+\boldsymbol{A}^\top\boldsymbol{\lambda}) = \frac{1}{9}\begin{bmatrix} 4 \\ 1 \end{bmatrix}$

Example 7.6 (Quadratic Program)
Consider the quadratic program

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \frac{1}{2}\begin{bmatrix}x_1\\x_2\end{bmatrix}^\top \begin{bmatrix}2 & 1\\1 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} + \begin{bmatrix}5\\3\end{bmatrix}^\top \begin{bmatrix}x_1\\x_2\end{bmatrix} \tag{7.46}$$

$$\text{subject to} \quad \begin{bmatrix}1 & 0\\-1 & 0\\0 & 1\\0 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} \leqslant \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \tag{7.47}$$

of two variables. The program is also illustrated in Figure 7.4. The objective function is quadratic with a positive semidefinite matrix $\boldsymbol{Q}$, resulting in elliptical contour lines. The optimal value must lie in the shaded (feasible) region, and is indicated by the star.



The unconstrained problem (indicated by the contour lines) has a minimum on the right side (indicated by the circle).

The box constraints ($-1 \leq x_1 \leq 1$ and $-1 \leq x_2 \leq 1$) require that the optimal solution is within the box, resulting in an optimal value indicated by the star.