Continuous Optimisation, Chpt 3: Constrained Convex Optimisation

Peter J.C. Dickinson
DMMP, University of Twente
p.j.c.dickinson@utwente.nl

http://dickinson.website/Teaching/2017CO.html

version: 06/10/17

Monday 2nd October 2017
**Theorem 3.1**

*Unconstrained quadratic optimisation can be solved analytically:*

\[ Q \not\preceq O \lor c \notin \text{Range}(Q) : \min_x \{ x^T Q x + 2c^T x + \gamma \} = -\infty. \]

\[ Q \succeq O \land c \in \text{Range}(Q) : \arg\min_x \{ x^T Q x + 2c^T x + \gamma \} = \{ x : Q x = -c \}. \]

**Theorem 3.2**

*The following problem is NP-hard (in \( n \) variables, not fixed):*

- **Minimising quadratic function with linear inequality constraints,**
  
  e.g. \( \min_x \{ x^T Q x : x \in \mathbb{R}^n_+ \} \).

  *[Murty and Kabadi, 1987]*
Complexity

Theorem 3.1

Unconstrained quadratic optimisation can be solved analytically:
\[
\begin{align*}
\text{for } & \mathbf{Q} \not\preceq \mathbf{O} \lor \mathbf{c} \not\in \text{Range}(\mathbf{Q}): \quad & \min_{\mathbf{x}} \{\mathbf{x}^T \mathbf{Q} \mathbf{x} + 2\mathbf{c}^T \mathbf{x} + \gamma\} = -\infty, \\
\text{for } & \mathbf{Q} \succeq \mathbf{O} \land \mathbf{c} \in \text{Range}(\mathbf{Q}): \quad & \arg\min_{\mathbf{x}} \{\mathbf{x}^T \mathbf{Q} \mathbf{x} + 2\mathbf{c}^T \mathbf{x} + \gamma\} = \{\mathbf{x} : \mathbf{Q} \mathbf{x} = -\mathbf{c}\}.
\end{align*}
\]

Theorem 3.2

The following problems are NP-hard (in \(n\) variables, not fixed):
- Minimising quadratic function with linear inequality constraints, e.g. \(\min_{\mathbf{x}} \{\mathbf{x}^T \mathbf{Q} \mathbf{x} : \mathbf{x} \in \mathbb{R}_+^n\}\). \([\text{Murty and Kabadi, 1987}]\)
- Minimising unconstrained quartic polynomial, e.g. \(f(\mathbf{y}) = (\mathbf{y}^2)^T \mathbf{Q} (\mathbf{y}^2)\).
**Complexity**

**Theorem 3.1**

*Unconstrained quadratic optimisation can be solved analytically:*

\[
\begin{align*}
\text{Q} \not\preceq O \lor c \not\in \text{Range}(Q) : \quad & \min_x \{x^TQx + 2c^Tx + \gamma \} = -\infty. \\
\text{Q} \succeq O \land c \in \text{Range}(Q) : \quad & \arg\min_x \{x^TQx + 2c^Tx + \gamma \} = \{x : Qx = -c\}.
\end{align*}
\]

**Theorem 3.2**

*The following problems are NP-hard (in n variables, not fixed):*

- Minimising quadratic function with linear inequality constraints, e.g. \(\min_x \{x^TQx : x \in \mathbb{R}^n_+\}\). \cite{MurtyKabadi1987}
- Minimising unconstrained quartic polynomial, e.g. \(f(y) = (y^{o2})^TQ(y^{o2})\).
- Checking whether quartic polynomial is convex. \cite{Ahmadietal2013}
Complexity Bibliography

M.R. Garey and D.S. Johnson. 

K.G. Murty and S.N. Kabadi. 
Some NP-complete problems in quadratic and nonlinear programming. 

Amir Ali Ahmadi, Alex Olshevsky, Pablo A. Parrilo and John N. Tsitsiklis. 
NP-hardness of deciding convexity of quartic polynomials and related problems. 
Literature & Survey

For further reading:
- FKS: p.273 – 277

Survey: https://goo.gl/forms/eR2fJrK7nAP3vifz1
Table of Contents

1 Introduction

2 Convex Problems
   • Unconstrained Recap
   • Constrained
   • Explicit problem
   • Active Indices

3 Descent Directions

4 KKT conditions
Optimality conditions in unconstrained optimization

\[
\inf_x f(x) \quad \text{s.t.} \quad x \in C \\
\text{(U)}
\]

\( C \subseteq \mathbb{R}^n \) is open set, \( f : C \rightarrow \mathbb{R}, \ f \in C^1. \)

**Lemma 3.3**

Consider the following statements:

1. \( x_0 \) is a global minimum of \( (U) \).
2. \( x_0 \) is a local minimum of \( (U) \).
3. \( \nabla f(x_0) = 0. \)

In general we have \( (1) \Rightarrow (2) \Rightarrow (3) \).

If \( f \) convex then \( (1) \Leftrightarrow (2) \Leftrightarrow (3) \).
Optimality conditions in constrained optimization

\[ \inf_x f(x) \quad \text{s.t.} \quad x \in F \quad (P) \]

where \( F \subseteq \mathbb{R}^n \), \( f : F \to \mathbb{R} \), \( f \in C^1 \).

**Lemma 3.4**

Consider the following statements for \( x_0 \in F \):

1. \( x_0 \) is a global minimum of \( (P) \).
2. \( x_0 \) is a local minimum of \( (P) \).
3. \( \nexists h \in \mathbb{R}^n \) such that \( \nabla f(x_0)^T h < 0 \) and \( x_0 + \varepsilon h \in F \) for all \( \varepsilon > 0 \) small enough.
4. \( \nabla f(x_0)^T (x - x_0) \geq 0 \) for all \( x \in F \).

In general we have \((1) \Rightarrow (2) \Rightarrow (3) \) and \((3) \iff (4)\).

If \( f \) and \( F \) convex then \((1) \iff (2) \iff (3) \iff (4)\).

**Ex. 3.1** Prove Lemma 3.4.
Explicit problem

\[
\begin{align*}
\inf_x & \quad f(x) \\
\text{s.t.} & \quad g_j(x) \leq 0 \quad \text{for all } j = 1, \ldots, m \\
& \quad x \in C.
\end{align*}
\]

where \( C \subseteq \mathbb{R}^n \) is an open set, \( f, g_1, \ldots, g_m \in C^1 \),
\[
\mathcal{F} := \{ x \in C : g_j(x) \leq 0 \text{ for all } j = 1, \ldots, m \}
\]

**Definition 3.5**

*We refer to \((C)\) as a convex (minimisation) problem if \( C \subseteq \mathbb{R}^n \) is a convex set and \( f, g_1, \ldots, g_m \) are convex functions on \( C \).*

**Ex. 3.2** Show that if \((C)\) is convex problem then \( \mathcal{F} \) is convex set.
Active indices

**Definition 3.6**

For \( x \in \mathcal{F} \), define index \( i \in \{1, \ldots, m\} \) to be **active** if \( g_i(x) = 0 \).

**Active index set** \( J_x \subseteq \{1, \ldots, m\} \) is the set of all active indices.

**Remark 3.7**

If \( g_1, \ldots, g_m \in C^0 \) then only constraints corresponding to active indices play a role at local minima.

**Ex. 3.3** Show that if \( g_1, \ldots, g_m \) are convex functions and \( x, y \in \mathcal{F} \) then \( \nabla g_i(x)^T(y - x) \leq 0 \) for all \( i \in J_x \).
# Table of Contents

1. Introduction

2. Convex Problems

3. Descent Directions
   - Sandwiching the condition
   - Slater’s Condition
   - Conclusion so far

4. KKT conditions
For fixed $x_0 \in F$, when does $\exists \hat{\varepsilon} > 0$ s.t. $x_0 + \varepsilon h \in F$ for all $\varepsilon \in (0, \hat{\varepsilon}]$?

Equivalently: $\exists \hat{\varepsilon} > 0$ s.t. $g_i(x_0 + \varepsilon h) \leq 0$ for all $i$ and all $\varepsilon \in (0, \hat{\varepsilon}]$.

Consider arbitrary $i \in \{1 \ldots, m\}$.

$$g_i(x_0 + \varepsilon h) = g_i(x_0) + \varepsilon \nabla g_i(x_0)^T h + o(\varepsilon).$$

- If $i \notin J_{x_0}$, or $\nabla g_i(x_0)^T h < 0$, or $g_i$ affine and $\nabla g_i(x_0)^T h \leq 0$: $\exists \hat{\varepsilon} > 0$ such that $g_i(x_0 + \varepsilon h) \leq 0$ for all $\varepsilon \in (0, \hat{\varepsilon}]$.

- If $i \in J_{x_0}$ and $\nabla g_i(x_0)^T h > 0$: $\nexists \hat{\varepsilon} > 0$ such that $g_i(x_0 + \varepsilon h) \leq 0$ for all $\varepsilon \in (0, \hat{\varepsilon}]$.

- If $i \in J_{x_0}$ and $\nabla g_i(x_0)^T h = 0$ and $g_i$ nonaffine: Condition may or may not hold.

Therefore

$$\left\{ h \in \mathbb{R}^n : \begin{array}{l} \nabla g_i(x_0)^T h \leq 0 \quad \text{for all } i \in J_{x_0} \text{ s.t. } g_i \text{ affine} \\ \nabla g_i(x_0)^T h < 0 \quad \text{for all } i \in J_{x_0} \text{ s.t. } g_i \text{ nonaffine} \end{array} \right\}$$

$$\subseteq \{ h \in \mathbb{R}^n : \exists \hat{\varepsilon} > 0 \text{ s.t. } x_0 + \varepsilon h \in F \text{ for all } \varepsilon \in (0, \hat{\varepsilon}] \}$$

$$\subseteq \{ h \in \mathbb{R}^n : \nabla g_i(x_0)^T h \leq 0 \text{ for all } i \in J_{x_0} \}$$
Sandwiching the condition

Lemma 3.8

Consider the following statements for $x_0 \in \mathcal{F}$:

(3a) $\nexists h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_i(x_0)^T h \leq 0$ for all $i \in J_{x_0}$.

(3) $\nexists h \in \mathbb{R}^n$ such that $\nabla f(x_0)^T h < 0$ and $x_0 + \varepsilon h \in \mathcal{F}$ for all $\varepsilon > 0$ small enough.

(3b) $\nexists h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_i(x_0)^T h < 0$ for all $i \in J_{x_0}$ s.t. $g_i$ is nonaffine, $\nabla g_j(x_0)^T h \leq 0$ for all $j \in J_{x_0}$ s.t. $g_j$ is affine.

In general we have $(3a) \Rightarrow (3) \Rightarrow (3b)$.

When are these three statements equivalent?
Slater’s Condition

**Definition 3.9**

We say that \( z \in \mathbb{R}^n \) is a Slater point for \((C)\) if \( z \in \mathcal{F} \) and \( g_i(z) < 0 \) for all \( i \) such that \( g_i \) is not an affine function.

Say that Slater’s condition holds for \((C)\) if there’s a Slater point.

**Example**

Consider \( m = n = 2 \), \( g_1(x) = x_1^2 + x_2^2 - 4 \), \( g_2(x) = x_2 - 1 \):

Which of the following are Slater points?

\[
A = (-2, -2)^T \quad B = (-1.2, -1.6)^T
\]
\[
C = (-0.5, -0.5)^T \quad D = (1, 1)^T
\]
\[
E = (-0.5, 1.5)^T \quad F = (-\sqrt{3}, 1)^T
\]

www.shakeq.com, login code utwente118
Lemma 3.10

If $x_0 \in F$ and problem (C) is a convex problem with Slater’s condition holding then the following are equivalent:

(3a) $\not\exists \ h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_i(x_0)^T h \leq 0$ for all $j \in J_{x_0}$.

(3) $\not\exists \ h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$ and $x_0 + \varepsilon h \in F$ for all $\varepsilon > 0$ small enough.

(3b) $\not\exists \ h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_i(x_0)^T h < 0$ for all $i \in J_{x_0}$ s.t. $g_i$ is nonaffine, $\nabla g_j(x_0)^T h \leq 0$ for all $j \in J_{x_0}$ s.t. $g_j$ is affine.
Conclusion so far

Lemma 3.11

For $x_0 \in F$ consider the following statements:

1. $x_0$ is a global minimiser of (C).
2. $x_0$ is a local minimiser of (C).
3. $\exists h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$ and $x_0 + \varepsilon h \in F$ for all $\varepsilon > 0$ small enough.
4. $\forall h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_j(x_0)^T h \leq 0$ for all $j \in J_{x_0}$.

In general we have $(1) \Rightarrow (2) \Rightarrow (3) \Leftarrow (4)$.

If (C) is convex then $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftarrow (4)$.

If (C) is convex and Slater’s condition holds then $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4)$.
<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

**1. Introduction**

**2. Convex Problems**

**3. Descent Directions**

**4. KKT conditions**

- Introduction
- Farkas’ Lemma
- Examples
- Example 1
- More KKT examples
We will now consider KKT conditions.

These are important conditions for optimality in constrained convex optimisation problems.

Originally called Kuhn-Tucker (KT) conditions after Harold W. Kuhn and Albert W. Tucker, who first published the conditions in 1951.

Conditions later found in the 1939 Master thesis of William Karush, hence now called Karush-Kuhn-Tucker (KKT) conditions.

“Searchers find, Researchers refind.”
Farkas’ Lemma

Lemma 3.12 (Farkas’ lemma)

Consider \( c, a_1, \ldots, a_m \in \mathbb{R}^n \). Then exactly one of the following statements holds:

1. \( \exists x \text{ s.t. } c^T x < 0 \text{ and } a_i^T x \leq 0 \text{ for all } i = 1, \ldots, m. \)
2. \( \exists y \in \mathbb{R}_+^m \text{ s.t. } c = -\sum_{i=1}^m y_i a_i. \)

https://ggbm.at/YadBMsQS

Corollary 3.13

For \( x_0 \in \mathcal{F} \) the following are equivalent:

1. \( \not\exists h \in \mathbb{R}^n \text{ s.t. } \nabla f(x_0)^T h < 0, \nabla g_i(x_0)^T h \leq 0 \text{ for all } i \in J_{x_0} \)
2. \( \nabla f(x_0) = -\sum_{i \in J_{x_0}} \lambda_i \nabla g_i(x_0), \text{ for some } \lambda_i \geq 0 \text{ } \forall i \in J_{x_0} \)
3. \( \exists \lambda \in \mathbb{R}_+^m \text{ s.t. } \lambda_i g_i(x_0) = 0 \forall i \text{ and } \nabla f(x_0) = -\sum_{i=1}^m \lambda_i \nabla g_i(x_0). \)

The last two conditions are referred to as Karush-Kuhn-Tucker (KKT) conditions, i.e. we say that the KKT conditions hold at such an \( x_0 \).
Conclusion

**Theorem 3.14**

For $x_0 \in \mathcal{F}$ consider the following statements:

1. $x_0$ is a global minimiser of (C).
2. $x_0$ is a local minimiser of (C).
3. $\nexists h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$ and $x_0 + \varepsilon h \in \mathcal{F}$ for all $\varepsilon > 0$ small enough.
4. $\nexists h \in \mathbb{R}^n$ s.t. $\nabla f(x_0)^T h < 0$, $\nabla g_i(x_0)^T h \leq 0$ for all $j \in \mathcal{J}_{x_0}$.
5. $\exists \lambda \in \mathbb{R}^m_+$ s.t. $\lambda_i g_i(x_0) = 0 \ \forall i$ and $\nabla f(x_0) = -\sum_{i=1}^{m} \lambda_i \nabla g_i(x_0)$.

In general we have $(1) \Rightarrow (2) \Rightarrow (3) \Leftrightarrow (4) \Leftrightarrow (5)$.

If (C) is convex then $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4) \Leftrightarrow (5)$.

If (C) is convex and Slater’s condition holds then $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4) \Leftrightarrow (5)$.
General idea to solving with KKT conditions

1. Transform to: \( \min_x \{ f(x) : g_i(x) \leq 0 \text{ for all } i = 1, \ldots, m \} \).

2. Analyse the problem: If the problem is **convex** and ...

   ...Slater’s cond. holds: KKT point \( \iff \) Global minimiser.

   ...Slater’s cond. doesn’t hold: KKT point \( \nRightarrow \) Global minimiser.

3. Find Solns \((x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m\) to following system of \(n + m\) equalities:

   \[
   \nabla f(x) = - \sum_{i=1}^{m} \lambda_i \nabla g_i(x),
   \lambda_i g_i(x) = 0 \quad \text{for all } i = 1, \ldots, m.
   \]

4. Solns with \(g_i(x) \leq 0 \forall i = 1, \ldots, m\) and \(\lambda \in \mathbb{R}_+^m\) are KKT points.

[Can combine points 3 and 4.]
Example 1: Steps 1 & 2: Transform & Analyse problem

For parameter $a \in \mathbb{R}$ consider problem

$$\min_x \{2x_1 + ax_2 : x_2 \leq 1, \ x_1^2 \leq x_2\}.$$

$$f(x) = 2x_1 + ax_2, \quad \nabla f = \begin{pmatrix} 2 \\ a \end{pmatrix}, \quad \nabla^2 f = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \text{Affine function}$$

$$g_1(x) = x_2 - 1, \quad \nabla g_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla^2 g_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \text{Affine function}$$

$$g_2(x) = x_1^2 - x_2, \quad \nabla g_2 = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}, \quad \nabla^2 g_2 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \text{Convex function}$$

Therefore convex minimisation problem.

$g_2$ is the only nonaffine constraint function, and for $x = (0, 1) \in \mathcal{F}$ have $g_2(x) = -1 < 0$. Therefore Slater’s condition holds.

Therefore a point is global minimiser if and only if it is KKT point.
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\nabla f(x) = - \sum_{i=1}^{2} \lambda_i \nabla g_i(x) \tag{1}
\]

\[
0 = \lambda_1 g_1(x) \tag{2}
\]

\[
0 = \lambda_2 g_2(x) \tag{3}
\]
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\begin{pmatrix}
2 \\
a
\end{pmatrix} = \nabla f(x) = - \sum_{i=1}^{2} \lambda_i \nabla g_i(x) = - \lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}
\]  

(1)

\[
0 = \lambda_1 g_1(x) = \lambda_1 (x_2 - 1) 
\]  

(2)

\[
0 = \lambda_2 g_2(x) = \lambda_2 (x_1^2 - x_2) 
\]  

(3)
Example 1: Steps 3: Solve equality requirements

Find $\mathbf{x} \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}^2$ such that

\[
\begin{pmatrix}
2 \\
a
\end{pmatrix} = \nabla f(\mathbf{x}) = -\sum_{i=1}^{2} \lambda_i \nabla g_i(\mathbf{x}) = -\lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} \tag{1}
\]

\[
0 = \lambda_1 g_1(\mathbf{x}) = \lambda_1(x_2 - 1) \tag{2}
\]

\[
0 = \lambda_2 g_2(\mathbf{x}) = \lambda_2(x_1^2 - x_2) \tag{3}
\]

Solution to (1) $\iff \lambda_2 \neq 0$ and $x_1 = -\lambda_2^{-1}$ and $\lambda_1 = \lambda_2 - a$.

As require $\lambda_2 \geq 0$, can restrict to $\lambda_2 > 0 > x_1$.

As $\lambda_2 > 0$, solution to (3) $\iff x_2 = x_1^2 = \lambda_2^{-2}$.

Two possible ways to satisfy (2):

- $\lambda_1 = 0$: Then $\mathbf{x} = (??, ??)^T$ and $\lambda = (0, ??)^T$.
- $x_2 = 1$: Then $\mathbf{x} = (??, 1)^T$ and $\lambda = (??, ??)^T$. 
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\begin{pmatrix}
2 \\
a
\end{pmatrix}
= \nabla f(x) = - \sum_{i=1}^{2} \lambda_i \nabla g_i(x) = - \lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}
\] (1)

\[ 0 = \lambda_1 g_1(x) = \lambda_1 (x_2 - 1) \] (2)

\[ 0 = \lambda_2 g_2(x) = \lambda_2 (x_1^2 - x_2) \] (3)

Solution to (1) \( \iff \lambda_2 \neq 0 \) and \( x_1 = -\lambda_2^{-1} \) and \( \lambda_1 = \lambda_2 - a \).

As require \( \lambda_2 \geq 0 \), can restrict to \( \lambda_2 > 0 > x_1 \).

As \( \lambda_2 > 0 \), solution to (3) \( \iff x_2 = x_1^2 = \lambda_2^{-2} \).

Two possible ways to satisfy (2):

- \( \lambda_1 = 0 \): Then \( x = (??, ??)^T \) and \( \lambda = (0, ??)^T \).
- \( x_2 = 1 \): Then \( x = (??, 1)^T \) and \( \lambda = (??, ??)^T \).
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\begin{bmatrix}
2 \\
a
\end{bmatrix} = \nabla f(x) = -\sum_{i=1}^{2} \lambda_i \nabla g_i(x) = -\lambda_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix}
\]

(1)

\[
0 = \lambda_1 g_1(x) = \lambda_1 (x_2 - 1)
\]

(2)

\[
0 = \lambda_2 g_2(x) = \lambda_2 (x_1^2 - x_2)
\]

(3)

Solution to (1) \( \Leftrightarrow \lambda_2 \neq 0 \) and \( x_1 = -\lambda_2^{-1} \) and \( \lambda_1 = \lambda_2 - a \).

As require \( \lambda_2 \geq 0 \), can restrict to \( \lambda_2 > 0 > x_1 \).

As \( \lambda_2 > 0 \), solution to (3) \( \Leftrightarrow x_2 = x_1^2 = \lambda_2^{-2} \).

Two possible ways to satisfy (2):

- \( \lambda_1 = 0 \): Then \( x = (??, ??)^T \) and \( \lambda = (0, a)^T \).

- \( x_2 = 1 \): Then \( x = (??, 1)^T \) and \( \lambda = (??, ??)^T \).
Example 1: Steps 3: Solve equality requirements

Find $\mathbf{x} \in \mathbb{R}^2$ and $\mathbf{\lambda} \in \mathbb{R}^2$ such that

$$
\begin{pmatrix} 2 \\ a \end{pmatrix} = \nabla f(\mathbf{x}) = - \sum_{i=1}^{2} \lambda_i \nabla g_i(\mathbf{x}) = - \lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix}
$$

(1)

$$0 = \lambda_1 g_1(\mathbf{x}) = \lambda_1 (x_2 - 1)
$$

(2)

$$0 = \lambda_2 g_2(\mathbf{x}) = \lambda_2 (x_1^2 - x_2)
$$

(3)

Solution to (1) $\Leftrightarrow \lambda_2 \neq 0$ and $x_1 = -\lambda_2^{-1}$ and $\lambda_1 = \lambda_2 - a$. As require $\lambda_2 \geq 0$, can restrict to $\lambda_2 > 0 > x_1$. As $\lambda_2 > 0$, solution to (3) $\Leftrightarrow x_2 = x_1^2 = \lambda_2^{-2}$.

Two possible ways to satisfy (2):

- $\lambda_1 = 0$: Then $\mathbf{x} = \begin{pmatrix} -a^{-1} \\ a^{-2} \end{pmatrix}^T$ and $\mathbf{\lambda} = \begin{pmatrix} 0 \\ a \end{pmatrix}^T$.
- $x_2 = 1$: Then $\mathbf{x} = \begin{pmatrix} ?? \\ 1 \end{pmatrix}^T$ and $\mathbf{\lambda} = \begin{pmatrix} ?? \\ ?? \end{pmatrix}^T$. 
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\begin{pmatrix} 2 \\ a \end{pmatrix} = \nabla f(x) = -\sum_{i=1}^{2} \lambda_i \nabla g_i(x) = -\lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} \tag{1}
\]

\[
0 = \lambda_1 g_1(x) = \lambda_1 (x_2 - 1) \tag{2}
\]

\[
0 = \lambda_2 g_2(x) = \lambda_2 (x_1^2 - x_2) \tag{3}
\]

Solution to (1) \( \iff \lambda_2 \neq 0 \) and \( x_1 = -\lambda_2^{-1} \) and \( \lambda_1 = \lambda_2 - a \).

As require \( \lambda_2 \geq 0 \), can restrict to \( \lambda_2 > 0 > x_1 \).

As \( \lambda_2 > 0 \), solution to (3) \( \iff x_2 = x_1^2 = \lambda_2^{-2} \).

Two possible ways to satisfy (2):

- \( \lambda_1 = 0 \): Then \( x = (-a^{-1}, a^{-2})^T \) and \( \lambda = (0, a)^T \).
- \( x_2 = 1 \): Then \( x = (-1, 1)^T \) and \( \lambda = (??, 1)^T \).
Example 1: Steps 3: Solve equality requirements

Find \( x \in \mathbb{R}^2 \) and \( \lambda \in \mathbb{R}^2 \) such that

\[
\begin{pmatrix} 2 \\ a \end{pmatrix} = \nabla f(x) = -\sum_{i=1}^{2} \lambda_i \nabla g_i(x) = -\lambda_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \lambda_2 \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix} \tag{1}
\]

\[0 = \lambda_1 \, g_1(x) = \lambda_1 (x_2 - 1) \tag{2}\]

\[0 = \lambda_2 \, g_2(x) = \lambda_2 (x_1^2 - x_2) \tag{3}\]

Solution to (1) \( \iff \lambda_2 \neq 0 \) and \( x_1 = -\lambda_2^{-1} \) and \( \lambda_1 = \lambda_2 - a \).

As require \( \lambda_2 \geq 0 \), can restrict to \( \lambda_2 > 0 > x_1 \).

As \( \lambda_2 > 0 \), solution to (3) \( \iff x_2 = x_1^2 = \lambda_2^{-2} \).

Two possible ways to satisfy (2):

- \( \lambda_1 = 0 \): Then \( x = (-a^{-1}, \ a^{-2})^T \) and \( \lambda = (0, \ a)^T \).
- \( x_2 = 1 \): Then \( x = (-1, \ 1)^T \) and \( \lambda = (1 - a, \ 1)^T \).
Example 1: Steps 4: Check inequality requirements

Have 2 candidate solutions, which fulfill all the KKT equality requirements. Will now check the inequality requirements:

1. \( \mathbf{x} = (-a^{-1}, a^{-2})^T \) and \( \mathbf{\lambda} = (0, a)^T \):
   - Require \( a \neq 0 \), otherwise solution ill defined.
   - Have \( \mathbf{\lambda} \in \mathbb{R}^2_+ \iff a \geq 0 \).
   - \( g_2(\mathbf{x}) = x_1^2 - x_2 = 0 \).
   - \( g_1(\mathbf{x}) = x_2 - 1 = a^{-2} - 1 \). Have \( g_1(\mathbf{x}) \leq 0 \iff |a| \geq 1 \).
     Inequalities satisfied if and only if \( a \geq 1 \).

2. \( \mathbf{x} = (-1, 1)^T \) and \( \mathbf{\lambda} = (1 - a, 1)^T \):
   - Have \( \mathbf{\lambda} \in \mathbb{R}^2_+ \iff a \leq 1 \).
   - \( g_1(\mathbf{x}) = x_2 - 1 = 0 \) and \( g_2(\mathbf{x}) = x_1^2 - x_2 = 0 \).
   - Inequalities satisfied if and only if \( a \leq 1 \).

N.B. If \( a = 1 \) for both have \( \mathbf{x} = (-1, 1)^T \) and \( \mathbf{\lambda} = (0, 1)^T \).
Example 1: Conclusion

If $a \leq 1$ have unique global minimiser at $x^* = (-1, 1)^T$.
This corresponds to $\lambda^* = (1 - a, 1)^T$.
Both constraints are active.
Optimal value equals $f(x^*) = -2 + a$.

If $a > 1$ have unique global minimiser at $x^* = (-a^{-1}, a^{-2})^T$.
This corresponds to $\lambda^* = (0, a)^T$.
Only the second constraint is active.
Optimal value equals $f(x^*) = -2a^{-1} + aa^{-2} = -a^{-1}$. 
KKT examples

Example
\[ \min_x \{ x^2 : (x - 2)^2 \leq 1 \} \]

Example
\[ \min_x \{ x^2 : (x - 2)^2 \leq 9 \} \quad \text{(can have } \mathcal{J}_x = \emptyset \text{ at minimum)} \]

Example
\[ \min_x \{ x^2 : (x - 2)^2 \leq 0 \} \quad \text{(KKT conditions not necessary)} \]

Example
For \( \mathbf{c}, \mathbf{a} \in \mathbb{R}^n \) with \( \mathbf{c} \neq \mathbf{0} \) consider the problem
\[ \min_x \{ \mathbf{c}^T \mathbf{x} : \| \mathbf{x} - \mathbf{a} \|^2 \leq 1 \} \]
Ex. 3.4 For each of the following problems, answer the following questions:

(a) Is the problem convex?
(b) Does Slater’s condition hold?
(c) What are the KKT points for this problem?
(d) What is the global minimiser to this problem?

1. \( \min_{x} \{ x_1 - x_2 : x_1^2 + x_2^2 \leq 4, \ x_2 \leq 1 \} \);
2. \( \min_{x} \{ x_1 : x_1^2 \leq x_2, \ x_2 \leq 0 \} \);
3. \( \min_{x} \{ x_1 : x_1 + x_2^2 \geq 1, \ x_1^3 \geq 0 \} \);
4. \( \min_{x} \{ c^T x : (x - a)^T Q (x - a) \leq 1 \} \), where \( Q \succ 0 \) and \( c \neq 0 \).