Exercise 8

January 20, 2021

1. Consider the univariate function:

   $f(x) = x^3 + 6x^2 - 3x - 5$

   Find its stationary points and indicate whether they are maximum, minimum, or saddle points.

   Solution:

   The first and the second derivatives of function 'f(x)' are given by

   $f'(x) = 3x^2 + 12x - 3,$
   $f''(x) = 6x$

   To find the positions of maxima and minima we solve equation for the points where the first derivative takes zero values:

   $3x^2 + 12x - 3 = 0 \Rightarrow x_{1,2} = -2 \pm \sqrt{5}$

   Substituting each of the roots into the second derivative we obtain that

   $f''(x_1) = f''(-2 + \sqrt{5}) > 0,$
   $f''(x_2) = f''(-2 - \sqrt{5}) < 0,$

   i.e. $x_1 = -2 + \sqrt{5}$ is the minimum, whereas $x_2 = -2 - \sqrt{5}$ is the maximum of function $f(x)$.

2. Express the following optimization problem as a standard linear pro-

gram in matrix notation
$\max_{x\in R^2,\xi\in R} p^T x + \xi$

Solution:

We first introduce vectors $\mathbf{y} = (x_0, x_1, \xi)^T$ and $\mathbf{c} = (p_0, p_1, 1)^T$, so that $\mathbf{p}^T\mathbf{x} + \xi = \mathbf{c}^T\mathbf{y}$.

We then introduce vector $\mathbf{b} = (0, 3, 0)$ and matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

which allow us to write the three constraints as $\mathbf{Ay} \leq \mathbf{b}$.

We now can write the problem as a standard linear program:

$\max_{\mathbf{y}\in R^3} \mathbf{c}^T\mathbf{y}$
subject to $\mathbf{Ay} \leq \mathbf{b}$

3. Consider the linear program:

$$\min_{x\in R^2} - \begin{bmatrix} 5 \\ 3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ subject to } \begin{bmatrix} 2 & 2 \\ 2 & -4 \\ -2 & 1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 33 \\ 8 \\ 5 \\ -1 \\ 8 \end{bmatrix}$$

Derive the dual linear program using Lagrange duality

Solution:

Let us define $\mathbf{c} = (-5, -3)^T$, $\mathbf{b} = (33, 8, 5, -1, 8)^T$ and

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & -4 \\ -2 & 1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix}$$

The linear program is then written as

$\min_{\mathbf{x}\in R^2} \mathbf{c}^T\mathbf{x}$
subject to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$

The Lagrangian of this problem is

$\mathcal{L}(\mathbf{x},\lambda) = \mathbf{c}^T\mathbf{x}+\lambda^T(\mathbf{A}\mathbf{x}-\mathbf{b}) = (\mathbf{c}^T\mathbf{x}+\lambda^T\mathbf{A})\mathbf{x}-\lambda^T\mathbf{b} = (\mathbf{c}\mathbf{x}+\mathbf{A}^T\lambda)^T\mathbf{x}-\lambda^T\mathbf{b}$

Taking gradient in respect to $\mathbf{x}$ and setting it to zero we obtain the extremum condition

$\mathbf{c}\mathbf{x} + \mathbf{A}^T\lambda = 0,$

that is

$\mathcal{D}(\lambda) = \min_{\mathbf{x}\in R^2} \mathcal{L}(\mathbf{x},\lambda) = -\lambda^T\mathbf{b}$

that is the dual problem is given by

$\max_{\lambda\in R^5} -\mathbf{b}^T\lambda$
subject to $\mathbf{c}\mathbf{x} + \mathbf{A}^T\lambda = 0$ and $\lambda \geq 0$

In terms of the original values of the parameters it can be thus written as

$$\max_{\lambda\in R^5} - \begin{bmatrix} 33 \\ 8 \\ 5 \\ -1 \\ 8 \end{bmatrix}^T \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{bmatrix}$$

$$\text{subject to } -\begin{bmatrix}5\\3\end{bmatrix}+\begin{bmatrix}2 & 2 & -2 & 0 & 0\\2 & -4 & 1 & -1 & 1\end{bmatrix}\begin{bmatrix}\lambda_1\\\lambda_2\\\lambda_3\\\lambda_4\\\lambda_5\end{bmatrix}=0$$

$$\text{and } \begin{bmatrix}\lambda_1\\\lambda_2\\\lambda_3\\\lambda_4\\\lambda_5\end{bmatrix}\geq 0$$

4. Consider the following convex optimization problem:

   $\min\limits_{w\in R^D} \frac{1}{2}w^T w$ subject to $w^T x \geq 1$

   Derive the Lagrangian dual by introducing the Lagrange multiplier λ.

   Solution:

   The primal problem can be written in teh standard form as

   $\min_{\mathbf{w}\in R^D} \frac{1}{2}\mathbf{w}^T\mathbf{w}$
   subject to $1-\mathbf{x}^T\mathbf{w}\leq 0$

   The Lagrangian is then

   $\mathcal{L}(\mathbf{w},\lambda)=\frac{1}{2}\mathbf{w}^T\mathbf{w}+\lambda(1-\mathbf{x}^T\mathbf{w})$

   Taking gradient in respect to $\mathbf{w}$ we obtain the position of the minimum: $\mathbf{w}=\lambda\mathbf{x}$

   Thus, the dual Lagrangian is

   $\mathcal{D}(\lambda)=\min_{\mathbf{w}\in R^D}\mathcal{L}(\mathbf{w},\lambda)=-\frac{\lambda^2}{2}\mathbf{x}^T\mathbf{x}+\lambda$

5. Consider the quadratic program:

   $$\min_{x\in R^2} \frac{1}{2}\begin{bmatrix}x_1\\x_2\end{bmatrix}^T\begin{bmatrix}2 & 1\\1 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}5\\3\end{bmatrix}^T\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

subject to $\begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Derive the dual quadratic program using Lagrange duality.

Solution:

We introduce $\mathbf{Q} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$,

$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Then the quadratic problem takes form

$\min_{\mathbf{x} \in R^2} \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x}$
subject to $\mathbf{A}\mathbf{x} \leq \mathbf{b}$

The Lagrangian corresponding to this problem is

$\mathcal{L}(\mathbf{x}, \lambda) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{c}^T\mathbf{x} + \lambda^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + (\mathbf{c}^T + \mathbf{A}^T\lambda)^T\mathbf{x} - \lambda^T\mathbf{b}$

where $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix}$ We minimize the Lagrangian by setting its gradient to zero, which results in

$\mathbf{Q}\mathbf{x} + \mathbf{c} + \mathbf{A}^T\lambda = 0 \Rightarrow \mathbf{x} = -\mathbf{Q}^{-1}(\mathbf{c} + \mathbf{A}^T\lambda)$

Substituting this back into the Lagrangian we obtain

$\mathcal{D}(\lambda) = \min_{\mathbf{x} \in R^2} \mathcal{L}(\mathbf{x}, \lambda) = -(\mathbf{c} + \mathbf{A}^T\lambda)^T\mathbf{Q}^{-1}(\mathbf{c} + \mathbf{A}^T\lambda) - \lambda^T\mathbf{b}$

The dual problem is now

$\max_{\lambda \in R^4} -(\mathbf{c} + \mathbf{A}^T\lambda)^T\mathbf{Q}^{-1}(\mathbf{c} + \mathbf{A}^T\lambda) - \lambda^T\mathbf{b}$
subject to $\lambda \geq 0$, where the parameter vectors and matrices are defined above.