Master of Technology

U2/6: Computational Intelligence I

To clarify slide 43 and 44

An easy optimization case

Consider the following constraint optimization problem to minimize

$$\min_{w}\left(\frac{1}{2}\|\mathbf{w}\|^2\right), \quad \text{subject to } g_i(w)=0, i=1,2,\ldots,l. \tag{1}$$

We can apply the method of Lagrange multipliers to solve it. In this method, we define the Lagrangian to be

$$L(w,\alpha)=\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\alpha_i g_i(w). \tag{2}$$

Here, the α_i are called the *Lagrange multipliers*. We would then find and set partial derivatives of L to be zero as

$$\frac{\partial L}{\partial w}=0, \frac{\partial L}{\partial \alpha_i}=0. \tag{3}$$

A difficult optimization case

Consider the following constraint optimization problem to minimize

$$\min_{w}\left(\frac{1}{2}\|\mathbf{w}\|^2\right), \quad \text{subject to} \quad g_i(w)\geq 0, i=1,2,\ldots,l. \tag{4}$$

Again, we define the Lagrangian to be below with all $\alpha_i \geq 0$

$$L(w,\alpha)=\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\alpha_i g_i(w). \tag{5}$$

We further define a function θ_p (p stands for *primal*)

$$\theta_p(w)=\max_{\alpha_i\geq 0}L(w,\alpha)=\max_{\alpha_i\geq 0}\left(\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\alpha_i g_i(w)\right). \tag{6}$$

Detailed study of (6)

- If the constraints are satisfied for a particular value of w, that means, $g_i(w)\geq 0$, then

$$\theta_p(w)=\max_{\alpha_i\geq 0}L(w,\alpha)=\max_{\alpha_i\geq 0}\left(\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\alpha_i \overbrace{g_i(w)}^{\text{positive}}\right)=\frac{1}{2}\|\mathbf{w}\|^2. \tag{7}$$

- If the constraints are NOT satisfied for a particular value of w, that means, $g_i(w)<0$, then

$$\theta_p(w)=\max_{\alpha_i\geq 0}L(w,\alpha)=\max_{\alpha_i\geq 0}\left(\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\alpha_i \overbrace{g_i(w)}^{\text{negative}}\right)=\infty. \tag{8}$$

Trick 1: Lagrangian function

Original optimization problem

Summarizing (7) and (8), we have

$$\min_{w}\left(\frac{1}{2}\|\mathbf{w}\|^2\right), \quad \text{subject to} \quad g_i(w) \geq 0, i = 1, 2, \ldots, l. \tag{9}$$

is as same as

$$\min_{w}\theta_p(w) = \min_{w}\max_{\alpha_i\geq 0} L(w,\alpha) = \min_{w}\max_{\alpha_i\geq 0}\left(\frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{l}\alpha_i g_i(w)\right). \tag{10}$$

Revised optimization problem

Trick 2: Dual representation

Revised optimization problem

Further revised optimization problem

$$\min_{w}\max_{\alpha_i\geq 0} L(w,\alpha) \geq \max_{\alpha_i\geq 0}\min_{w} L(w,\alpha),$$

$$\min_{w}\overbrace{\theta_p(w)}^{primal} \geq \max_{\alpha_i\geq 0}\overbrace{\theta_d(w)}^{dual},$$