تمرین شماره دوم درس یادگیری ماشین

پریا پاسه‌ورز

شماره دانشجویی: 810101393

سوال ششم)

برای مسئله بهینه‌سازی زیر با استفاده از ضرایب لاگرانژ، دوگان مسئله را به دست آورید و آن را بر حسب ضرایب لاگرانژ بنویسید.

$$\min_{x \in R^{2}} \boldsymbol{f}(\boldsymbol{x}) = \frac{\mathbf{1}}{\mathbf{2}} \boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{c}^{\boldsymbol{T}} \boldsymbol{x}$$

$$\textbf{subject to } \mathbf{Ax} \leq \mathbf{b}$$

$$\mathbf{Q} \in \mathbf{R}^{2\times 2}, \mathbf{c} \in \mathbf{R}^{2}, \mathbf{A} \in \mathbf{R}^{3\times 2}, \mathbf{b} \in \mathbf{R}^{3}$$

رابطه تابع Lagrangian برای یک مسئله constrained به شکل زیر است:

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

که در اینجا f(x) مسئله‌ای که می‌خواهیم بهینه کنیم و g(x) در واقع constrained نامساوی مسئله است. حال با جایگذاری صورت مسئله خواهیم داشت:

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

حال باید برقراری شرایط Karush-Kuhn-Tucker (KKT) را بررسی کنیم:

اگر مسئله به شکل زیر باشد:

$$\min_{\mathrm{x} \in \mathrm{R}^{\mathrm{n}}} \mathrm{f}(\mathrm{x})$$

$$\text{subject to } \mathrm{h_i}(\mathrm{x}) \leq 0, \quad \mathrm{i} = 1, \dots, \mathrm{m}$$

$$\mathrm{l_j}(\mathrm{x}) = 0, \quad \mathrm{j} = 1, \dots, \mathrm{r}$$

- Stationary:

$$0 \in \partial f(x) + \sum_{i=1}^{m} u_i \partial h_i(x) + \sum_{j=1}^{r} v_j \partial l_j(x)$$

که در مسئله داده شده خواهیم داشت:

$$\nabla_x \mathcal{L}(x, \lambda) = Qx + c + A^T \lambda = 0$$
$$Qx + c + A^T \lambda = 0 \;\Rightarrow\; x = -Q^{-1}(c + A^T \lambda)$$

- Primal feasibility:

$$l_i(x) = 0, \quad h_i(x) \leq 0 \quad \forall i$$

این خاصیت اطمینان حاصل می‌کند که *x یافت شده باید در constrainedهای اولیه مسئله صدق کند، یعنی در اینجا باید در $\mathrm{Ax} \leq \mathrm{b}$ صدق کند.

- Dual feasibility:

$$\lambda_i \geq 0, \quad \forall i$$

این خاصیت اطمینان حاصل می‌کند ضرایب لاگرانژ نامنفی هستند:

$$\lambda \geq 0$$

- Complementary slackness:

$$u_i \cdot h_i(x) = 0 \text{ for all } i$$

که در مسئله داده شده خواهیم داشت:

$$\lambda_i\left(a_i^T x - b_i\right) = 0 \quad \forall i$$

حال با کمک ویژگی‌های به دست آمده، تابع دوگان Lagrangian را حساب می‌کنیم:

$$g(u,v) = \min_{x \in R^n} L\,(x,u,v)$$

باید x به دست آمده از خاصیت stationary را در $\mathcal{L}(x,\lambda)$ جایگذاری کنیم:

$$g(\lambda) = \mathcal{L}(-Q^{-1}(c + A^T\lambda), \lambda)$$

$$= \frac{1}{2}\left(-Q^{-1}(c + A^T\lambda)\right)^T Q\left(-Q^{-1}(c + A^T\lambda)\right) + c^T\left(-Q^{-1}(c + A^T\lambda)\right) + \lambda^T\left(A\left(-Q^{-1}(c + A^T\lambda)\right) - b\right)$$

$$= \frac{1}{2}(c + A^T\lambda)^T Q^{-1} Q Q^{-1}(c + A^T\lambda) - (c^T Q^{-1} c + c^T Q^{-1} A^T\lambda) - \lambda^T A Q^{-1}(c + A^T\lambda) - \lambda^T b$$

$$g(\lambda) = -\frac{1}{2}(c + A^T\lambda)^T Q^{-1}(c + A^T\lambda) - \lambda^T b$$

پس مسئله دوگان خواهد بود:

$$\max_{\lambda \geq 0} g\,(\lambda)$$

$$g(\lambda) = -\frac{1}{2}(c + A^T\lambda)^T Q^{-1}(c + A^T\lambda) - \lambda^T b$$