MML - Homework 3

Sunday, February 19, 2023 7:51 PM

1.1: $f(x) = \|x\|_2^2,\ h(x) = \sum_{i=1}^{n} x_i - 1.$

Want to find coupled equations for Lagrangian to be stationary w.r.t x and λ.

$L(x,\lambda) = \|x\|_2^2 + \lambda\left(\sum_{i=1}^{n} x_i - 1\right)$

$$\nabla_L = \frac{\partial L}{\partial x} = 2\|x\|_2 + \lambda\left(\sum_{i=1}^{n} x_i \cancel{-1}\right)$$

$$= 2\|x\|_2 + \lambda n$$

$$= 2\|x\|_2 + \lambda n = 0 \longrightarrow \lambda = -\frac{2\|x\|_2}{n}$$

$$\frac{\partial L}{\partial \lambda} = \cancel{\|x\|_2^2} + \lambda\sum_{i=1}^{n} x_i \cancel{-1}$$

$$= \sum_{i=1}^{n} x_i$$

$$= x_n \longrightarrow x_n = 0$$

1.2: $f(x) = \sum_{i=1}^{n} x_i,\ h(x) = \|x\|_2^2 - 1.$

$L(x,\lambda) = \sum_{i=1}^{n} x_i + \lambda\left(\|x\|_2^2 - 1\right)$

$$\nabla_L = \frac{\partial L}{\partial x} = n + 2\|x\|_2\lambda = 0 \longrightarrow \lambda = -\frac{n}{2\|x\|_2}$$

$$\frac{\partial L}{\partial \lambda} = \|x\|_2^2 \longrightarrow x_2 = 0$$

1.3: $f(x) = \|x\|_2^2,\ h(x) = x^T Q x - 1.$

$L(x,\lambda) = \|x\|_2^2 + \lambda\left(x^T Q x - 1\right)$

$$\nabla_L = \frac{\partial L}{\partial x} = 2\|x\|_2 + \lambda\left(x^T Q^T + x^T Q\right) = 0 \longrightarrow \lambda = -\frac{2\|x\|_2}{\left(x^T Q^T + x^T Q\right)}$$

$$\frac{\partial L}{\partial \lambda} = x^T Q x = 0 \longrightarrow x = 0$$

2: $L(x_1,\ldots,x_n,\lambda) = f(x_1,\ldots,x_n) - \lambda\left(g(x_1,\ldots,x_n) - 1\right)$

$$\frac{\partial L}{\partial x_i} = a_i x_i^{(a_i - 1)} + \lambda\prod_{j \neq i} x_j^{a_j} = 0$$

$$\frac{\partial L}{\partial \lambda} = \prod_{i=1}^{n} x_i^{(a_i)} - 1 = 0$$

After using the constraint, we find that $\lambda = -\sum_{i=1}^{n} \frac{a_i}{x_i}$ $\longrightarrow$ we then substitute this back into the first equation and ln both sides to get:

$$\sum_{i=1}^{n} a_i \ln(x_i) \leq \ln\left(\sum_{i=1}^{n} a_i x_i\right)$$

This proves the inequality due to natural log properties.

3: $\min_{x \to y} f(x,y) = \frac{1}{2}((x-x_0)^2 + (y-y_0)^2)$ with (x_0, y_0) fixed and subject to $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Lagrangian: $L(x,y,\lambda) = f(x,y) + \lambda((g(x,y) - c)$

$$L(x,y,\lambda) = \frac{1}{2}((x-x_0)^2 + (y-y_0)^2) + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right)$$

Solving for x, y: $\frac{\partial L}{\partial x} = (x - x_0) + \frac{2\lambda x}{a^2} = 0$

$$\frac{a^2x + 2\lambda x}{a^2} = x_0$$

$$\frac{x(a^2+2\lambda)}{a^2} = x_0$$

$$x = \frac{x_0 a^2}{(a^2+2\lambda)}$$

$$\frac{\partial L}{\partial y} = (y - y_0) + \frac{2\lambda y}{b^2} = 0$$

$$y = \frac{y_0 b^2}{(b^2+2\lambda)}$$

Rewriting ellipse constraint in terms of Lagrange parameter: $\frac{\partial L}{\partial \lambda} = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$

$$1 = \frac{\left(\frac{x_0 a^2}{(a^2+2\lambda)}\right)^2}{a^2} + \frac{\left(\frac{y_0 b^2}{(b^2+2\lambda)}\right)^2}{b^2} \cancel{-1}$$

$$1 = \frac{a^2+2\lambda}{a^2(x_0 a^2)} + \frac{b^2+2\lambda}{b^2(y_0 b^2)}$$

$$g(\lambda) = \frac{a^2+2\lambda}{a^4 x_0} + \frac{b^2+2\lambda}{b^4 y_0} - 1 = 0$$

Show it leads to quartic, can general quartic be solved with radicals?

$$(\,(a^2+2\lambda)b^4 y_0 + (b^2+2\lambda)a^4 x_0 - a^4 x_0 b^4 y_0\,) = 0$$

Expand and simplify: $\lambda^4(2a^2b^2x_0y_0) + \lambda^3(a^4y_0^2 + b^4x_0^2 - a^4x_0b^4) + \lambda^2(a^2b^4y_0 + b^2a^4x_0 - a^4x_0b^4) + 2\lambda(a^4x_0b^4) = 0$

A general quartic can't be solved by radicals, but in some cases when the quartic is symmetric, it can. (Normally you can find an approx. solution via special methods)

4: $E(\omega) = \|y - \sum_{k=1}^{K} w_k x_k\|_2^2$ where $w \in \mathbb{R}^K$ and $x_i, y \in \mathbb{R}^D$

Make Lagrangian with constraint $\sum_{k=1}^{K} \omega_k = 1$: $L(\omega,\lambda) = \|y - \sum_{k=1}^{K} \omega_k x_k\|_2^2 + \lambda(\sum_{k=1}^{K} \omega_k - 1)$

Solve for Lagrange parameter. Explain the steps: Take partial derivatives to ω_k and set equal to 0:

$$\frac{\partial L}{\partial \omega_i} = -2(x_i^T y) + 2(x_i^T \omega \underbrace{[x_1, x_2, \dots x_K]}_{\text{Represent as } X}) + \lambda = 0$$

Note: $E(\omega) = (y - \sum_{k=1}^{K} w_k x_k)^{\overset{\text{Transpose}}{T}} (y - \sum_{k=1}^{K} \omega_k x_k)$

Rewrite to matrix form as: $-2X^T y + 2X^T X\omega + \lambda 1_K = 0$

Solve for λ: $\lambda = \frac{2(X^T y - X^T X\omega)}{K}$

We are dividing element by element due to the constraint that the weights sum to 1.

Solve for ω while satisfying constraint: Substitute λ into Lagrangian:

$L(\omega,\lambda) = \|y - X\omega\|_2^2 + \lambda(1 - 1_K^T \omega)$

Take the gradient of $L(\omega,\lambda)$ w.r.t ω:

$\nabla L(\omega,\lambda) = -2X^T y + 2X^T X\omega - \lambda 1_K = 0$

Note: $1_K^T \omega = 1$.

We use the condition and multiply both sides by X^T, then invert $X^T X$.

$$\omega = (X^T X)^{-1} X^T (y - \frac{\lambda}{K} 1_K)$$

(Assuming $X^T X$ is invertible)

5. https://colab.research.google.com/drive/1yUwtI75cuRYvn-xzslW-cJ9x6Njm5y4x?usp=sharing