Reinfocement Learning

Dr M.Rohban

Department: Electrical Engineering

Amirreza Velae 400102222
Webpage amirrezavelae.github.io

Homework 1

March 6, 2024

Reinfocement Learning

Homework 1

Amirreza Velae 400102222
Webpage amirrezavelae.github.io

Optimization

Convexity

Prove that for every convex function $f(x)$, every critical point is a global minimum.

solution

We know that a function $f(x)$ is convex if and only if for every $x_1, x_2 \in \mathbb{R}$ and $0 \leq \lambda \leq 1$:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

Or by first order condition:

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$$

Now let's assume that x_1 is a critical point of $f(x)$, then $f'(x_1) = 0$. Now let's assume that x_2 is any other point in the domain of $f(x)$, then we have:

$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1) = f(x_1)$$

So we can conclude that x_1 is a global minimum of $f(x)$.
If $f(x)$ is strictly convex, then x_1 is the only global minimum of $f(x)$.

Convex Optimization

Sove the following optimization problem:

$$\begin{aligned} \min_{x} \quad & f(x,y,z) = x^2 + y^2 + z^2 \\ \text{s.t.} \quad & z^2 = x^2 + y^2 \\ & z = x + y + 1 \end{aligned}$$

solution

We can use Lagrange multipliers to solve this problem. We define the Lagrangian as:

$$L(x,y,z,\lambda) = f(x,y,z) - \lambda_1(z^2 - x^2 - y^2) - \lambda_2(z - x - y - 1)$$
$$L(x,y,z,\lambda) = x^2 + y^2 + z^2 - \lambda_1(z^2 - x^2 - y^2) - \lambda_2(z - x - y - 1)$$

Now we can find the critical points of $L(x, y, z, \lambda)$ by solving the following system of equations:

$$\frac{\partial L}{\partial x} = 2x + 2\lambda_1 x + \lambda_2 = 0$$
$$\frac{\partial L}{\partial y} = 2y + 2\lambda_1 y + \lambda_2 = 0$$
$$\frac{\partial L}{\partial z} = 2z - 2\lambda_1 z - \lambda_2 = 0$$
$$\frac{\partial L}{\partial \lambda_1} = z^2 - x^2 - y^2 = 0$$
$$\frac{\partial L}{\partial \lambda_2} = z - x - y - 1 = 0$$

From first and second equations we have:

$$2(x + \lambda_1) + \lambda_2 = 0$$
$$2(y + \lambda_1) + \lambda_2 = 0$$
$$\implies x + \lambda_1 = y + \lambda_1 \implies x = y$$

Now we can reduce the system of equations to:

$$2x + 2\lambda_1 x + \lambda_2 = 0$$
$$2z - 2\lambda_1 z - \lambda_2 = 0$$
$$z^2 - 2x^2 = 0$$
$$z - 2x - 1 = 0$$

Now we can solve the system of equations to find the critical points:

$$z = 2x + 1 \to z^2 = 4x^2 + 4x + 1$$
$$z^2 = 2x^2 \to 4x^2 + 4x + 1 = 2x^2 \to 2x^2 + 4x + 1 = 0$$
$$x = -1 + \frac{\sqrt{2}}{2} \quad \text{or} \quad x = -1 - \frac{\sqrt{2}}{2}$$
$$y = -1 + \frac{\sqrt{2}}{2} \quad \text{or} \quad y = -1 - \frac{\sqrt{2}}{2}$$
$$z = 2x + 1 \implies z = \sqrt{2} \quad \text{or} \quad z = -\sqrt{2}$$

So the critical points are:

$$(x, y, z) = \left(-1 + \frac{\sqrt{2}}{2}, -1 + \frac{\sqrt{2}}{2}, \sqrt{2}\right)$$
$$(x, y, z) = \left(-1 - \frac{\sqrt{2}}{2}, -1 - \frac{\sqrt{2}}{2}, -\sqrt{2}\right)$$

Now we can find the value of $f(x, y, z)$ at these points:

$$f\left(-1+\frac{\sqrt{2}}{2}, -1+\frac{\sqrt{2}}{2}, \sqrt{2}\right) = 2$$
$$f\left(-1-\frac{\sqrt{2}}{2}, -1-\frac{\sqrt{2}}{2}, -\sqrt{2}\right) = 2$$

So the minimum value of $f(x, y, z)$ is 2 and it is achieved at the critical points.

Duality

We define the dual problem of a convex optimization problem as:

$$\max_{\lambda,\mu} \quad L(\lambda, \mu) = \min_x \left\{f(x) + \lambda^T g(x) + \mu^T h(x)\right\}$$
$$L(\lambda, \mu) = \min_x \left\{f(x) + \lambda^T g(x) + \mu^T h(x)\right\}$$

Prove that :

$$L(\lambda, \mu) \leq min_x f(x) \quad \forall \lambda, \mu$$

and give an example of a convex optimization problem for which the equality does not hold for any λ, μ.

solution

We know that $f(x)$ is a convex function and $g(x)$ and $h(x)$ are affine functions. So the function $L(\lambda, \mu)$ is a concave function.
Now let's assume that x^* is the optimal solution of the primal problem, then we have:

$$f(x^*) + \lambda^T g(x^*) + \mu^T h(x^*) \geq f(x) + \lambda^T g(x) + \mu^T h(x) \quad \forall x$$

So we can conclude that:

$$f(x^*) + \lambda^T g(x^*) + \mu^T h(x^*) \geq \min_x \left\{f(x) + \lambda^T g(x) + \mu^T h(x)\right\}$$

So we can conclude that:

$$L(\lambda, \mu) \leq \min_x f(x)$$

Now let's consider the following optimization problem:

$$\begin{aligned} \min_x \quad & f(x) = x^2 \\ \text{s.t.} \quad & x \geq 1 \end{aligned}$$

The optimal solution of this problem is $x^* = 1$ and $f(x^*) = 1$. Now let's consider the dual problem:

$$\max_\lambda \quad L(\lambda) = \min_x \left\{x^2 + \lambda(x-1)\right\}$$

The optimal solution of this problem is $\lambda^* = 0$ and $L(\lambda^*) = 0$. So we can conclude that:

$$L(\lambda^*) = 0 < f(x^*) = 1$$

So the equality does not hold for any λ, μ.
The reason that the equality does not hold is that the constraint is not convex.

Water Filling

Solve the following optimization problem:

$$\begin{aligned} \max_{p} \quad & -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \text{s.t.} \quad & x \geq 0 \quad \sum_{i=1}^{n} x_i = 1 \end{aligned}$$

solution

We can use Lagrange multipliers to solve this problem. We define the Lagrangian as:

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

Now we can find the critical points of $L(x, \lambda)$ by solving the following system of equations:

$$\begin{aligned} \frac{\partial L}{\partial x_i} &= -\frac{1}{\alpha_i + x_i} - \lambda = 0 \quad \forall i \in \{1, 2, \ldots, n\} \\ \frac{\partial L}{\partial \lambda} &= \sum_{i=1}^{n} x_i - 1 = 0 \end{aligned}$$

From the first equation we have:

$$-\frac{1}{\alpha_i + x_i} - \lambda = 0 \to x_i = -\alpha_i - \frac{1}{\lambda}$$

Now we can use the second equation to find the value of λ:

$$\sum_{i=1}^{n} x_i = 1 \to \sum_{i=1}^{n} -\alpha_i - \frac{1}{\lambda} = 1 \to -\frac{1}{\lambda} = 1 - \sum_{i=1}^{n} \alpha_i \to \lambda = -\frac{1}{1 - \sum_{i=1}^{n} \alpha_i}$$

Now we can use the value of λ to find the value of x_i:

$$x_i = -\alpha_i - \frac{1}{\lambda} = -\alpha_i - \frac{1}{-\frac{1}{1-\sum_{i=1}^{n} \alpha_i}} = -\alpha_i + 1 - \sum_{i=1}^{n} \alpha_i = 1 - \alpha_i$$

So the optimal solution of the problem is:

$$x_i = 1 - \alpha_i \quad \lambda = -\frac{1}{1 - \sum_{i=1}^{n} \alpha_i}$$

Maximum Entropy

Solve the following optimization problem:

$$\begin{aligned} \max_{p} \quad & -\int p(x) \log p(x) \, \mathrm{d}x \\ \text{s.t.} \quad & \mathbb{E}_p[X] = \mu \quad \mathbb{V}_p[X^2] = \sigma^2 \end{aligned}$$

solution

Here $p(x)$ is the probability mass function of a random variable X. We know that probability mass function is a non-negative function and $\sum_x p(x) = 1$; i.e:

$$p(x) \geq 0 \quad \text{and} \quad \int p(x) \, \mathrm{d}x = 1$$

We can use variational calculus to solve this problem. We define the $L(p, \lambda)$ as:

$$L(p, \lambda) = -\int p(x) \log p(x) \, \mathrm{d}x - \lambda \left(\int p(x) \, \mathrm{d}x - 1 \right)$$

Now, using the variational calculus given in the homework, we know that the critical points of $\int F(x, y, y') \, \mathrm{d}x$ are the solutions of the following system of equations:

$$\frac{d}{dx} \frac{\partial F}{\partial y'} - \frac{\partial F}{\partial y} = 0$$

Now we can find the critical points of $L(p, \lambda)$ by solving the following system of equations:

$$\begin{aligned} \frac{d}{dx} \frac{\partial L}{\partial p'} - \frac{\partial L}{\partial p} = 0 \\ \implies \frac{d}{dx} \times 0 - \frac{\partial L}{\partial p} = 0 \end{aligned}$$

Now we can solve $\frac{\partial L}{\partial p} = 0$ to find the critical points of $L(p, \lambda)$:

$$\begin{aligned} \frac{\partial L}{\partial p} &= -\log p(x) - 1 - \lambda = 0 \\ \implies p(x) &= e^{-1-\lambda} \end{aligned}$$

Now we can use the constraints to find the value of λ:

$$\begin{aligned} &\int p(x) \, \mathrm{d}x = 1 \to \int e^{-1-\lambda} \, \mathrm{d}x = e^{-1-\lambda} \times (b - a) = 1 \\ &\implies e^{-1-\lambda} = \frac{1}{b-a} \to \lambda = \log(b-a) - 1 \\ &\implies p(x) = \frac{1}{b-a} \end{aligned}$$

So the probability mass function that maximizes the entropy is a uniform distribution. To find the value of μ and σ^2 we can use the following equations:

$$\begin{aligned}
\mathbb{E}_p[X] &= \int xp(x)\,\mathrm{d}x = \frac{1}{b-a}\int_a^b x\,\mathrm{d}x = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{b+a}{2} = \mu \\
\mathbb{V}_p[X] &= \int x^2 p(x)\,\mathrm{d}x - \mu^2 = \frac{1}{b-a}\int_a^b x^2\,\mathrm{d}x - \mu^2 \\
&= \frac{1}{b-a}\frac{b^3-a^3}{3} - \mu^2 = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} = \frac{a^2-ab+b^2}{12} = \sigma^2
\end{aligned}$$

So the optimal solution of the problem is:

$$p(x) = \frac{1}{b-a} \quad \mu = \frac{b+a}{2} \quad \sigma^2 = \frac{a^2-ab+b^2}{12}$$

Or:

$$\begin{cases}\mu = \frac{b+a}{2} \\ \sigma = \frac{|b-a|}{\sqrt{12}} = \frac{b-a}{2\sqrt{3}}\end{cases} \implies \begin{cases} a = \mu - \sqrt{3}\sigma \\ b = \mu + \sqrt{3}\sigma\end{cases} \implies p \sim U(\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma)$$

Note: If there was any limit on the domain of x; i.e. $\mathrm{dom}(x) = R$, then the optimal solution would be a gaussian distribution(Wikipedia).
As a short proof, we can use the following Lagrangian:

$$\begin{aligned}
L(p,\lambda,\mu,\sigma) &= -\int p(x)\log p(x)\,\mathrm{d}x - \lambda_0\left(\int p(x)\,\mathrm{d}x - 1\right) - \lambda_1\left(\int (x-\mu)^2 p(x)\,\mathrm{d}x - \sigma^2\right) \\
&= -\int p(x)(\log p(x) - \lambda_0 - \lambda_1(x-\mu)^2)\,\mathrm{d}x + \lambda_0 + \lambda_1\sigma^2
\end{aligned}$$

Using calculus of variations, we can find the critical points of $L(p,\lambda,\mu,\sigma)$ by solving the following system of equations:

$$\begin{aligned}
&\frac{d}{dx}\frac{\partial L}{\partial p'} - \frac{\partial L}{\partial p} = 0 \\
&\frac{d}{dx}\times 0 - \frac{\partial L}{\partial p} = 0 \\
&\frac{\partial L}{\partial p} = -\log p(x) - 1 - \lambda_0 - \lambda_1(x-\mu)^2 = 0 \\
&\implies p(x) = e^{-1-\lambda_0-\lambda_1(x-\mu)^2}
\end{aligned}$$

Now, using the constraints, we can find the value of λ_0 and λ_1:

$$\begin{aligned}
&\int p(x)\,\mathrm{d}x = 1 \to \int e^{-1-\lambda_0-\lambda_1(x-\mu)^2}\,\mathrm{d}x = 1 \\
&\implies e^{-1-\lambda_0}\int e^{-\lambda_1(x-\mu)^2}\,\mathrm{d}x = 1
\end{aligned}$$

We can use the following substitution to solve the integral:

$$t = \sqrt{\lambda_1}(x-\mu) \to \mathrm{d}t = \sqrt{\lambda_1}\,\mathrm{d}x$$

So we can conclude that:

$$\int e^{-\lambda_1(x-\mu)^2}\,\mathrm{d}x = \frac{1}{\sqrt{\lambda_1}}\int e^{-t^2}\,\mathrm{d}t = \frac{\sqrt{\pi}}{\sqrt{\lambda_1}}$$

So we can conclude that:

$$e^{-1-\lambda_0}\frac{\sqrt{\pi}}{\sqrt{\lambda_1}} = 1 \to e^{-1-\lambda_0} = \frac{\sqrt{\lambda_1}}{\sqrt{\pi}} \to \lambda_0 = -1 - \log\left(\sqrt{\pi}\right) - \frac{1}{2}\log(\lambda_1)$$

Now we can use the second constraint to find the value of λ_1:

$$\int (x-\mu)^2 p(x)\,\mathrm{d}x = \sigma^2 \to \int (x-\mu)^2 e^{-1-\lambda_0-\lambda_1(x-\mu)^2}\,\mathrm{d}x = \sigma^2$$

We can use the same substitution to solve the integral:

$$\frac{1}{\sqrt{\lambda_1}}\int \frac{t^2}{\lambda_1}e^{-1-\lambda_0-t^2}\,\mathrm{d}t = \frac{e^{-1-\lambda_0}}{\lambda_1\sqrt{\lambda_1}}\int t^2 e^{-t^2}\,\mathrm{d}t = \frac{\frac{\sqrt{\lambda_1}}{\sqrt{\pi}}}{\lambda_1\sqrt{\lambda_1}}\int t^2 e^{-t^2}\,\mathrm{d}t = \frac{1}{\lambda_1\sqrt{\pi}}\frac{\sqrt{\pi}}{2} = \sigma^2$$
$$\to \frac{1}{2\lambda_1} = \sigma^2 \to \lambda_1 = \frac{1}{2\sigma^2}$$

So the optimal solution of the problem is:

$$p(x) = e^{-1-\lambda_0-\lambda_1(x-\mu)^2} \quad \lambda_0 = -1 - \log\left(\sqrt{\pi}\right) - \frac{1}{2}\log(\lambda_1) \quad \lambda_1 = \frac{1}{2\sigma^2}$$

Or:

$$p(x) = e^{-1+\log\left(\sqrt{\pi}\right)+\frac{1}{2}\log\left(2\sigma^2\right)-\frac{1}{2\sigma^2}(x-\mu)^2} = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

So the optimal solution of the problem is a gaussian distribution.

End of Homework 1