# Derivation of determining extrema of a function with MaxEnt formulation

Ravi Sankar Saripalli

May 27, 2019

## Description

This is an attempt to see how determination of extrema of a function can be determined with MaxEnt formulation. Given $f(x)$ find its extrema (max/min). Usual way to achieve this is to manipulate $x$ to determine extrema looking for $x$ where function derivative $f'(x)$ vanishes. An alternate approach is to use probability distribution function $p(x)$ as an independant parameter and seek to maximize expected value of $f(x)$ with the condition that entropy (as defined by Shanon) of the probability distribution $p(x)$ is maximized, thus ensuring that the distribution function derived is least biased.

## The MaxEnt Lagrangian

$$\boxed{\mathcal{L}(p,\lambda) = \int f(x)p(x)dx + T\int p(x)ln(p(x))dx + \lambda\left\{\left(\int p(x)dx\right) - 1\right\}} \tag{1}$$

The first term in the Lagrangian is the expected value of the function corresponding to the PDF $p(x)$, the second term corresponds to the entropy of the PDF scaled by an arbitray constant $T$ and the last term corresponds to the equality constraint that integral of PDF is one. The multiplier $\lambda$ is the Lagrangian parameter.

Noting that the Lagrangian is function of $p(x)$ and $\lambda$, the extrema of the Lagrangian is obtained by requiring the functional derivative with respect $p(x)$ and the Lagrangian parameter $\lambda$ are zero.

Although this enables us to determine $p(x)$ and $\lambda$ corresponding to the extrema, we need additional criterion to determine if the extrema is either minimum or maximum. Whether one can use the sign of the second derivative in functional space similar to simple variables can be used for this purpose is yet to be explored by me....

## The Euler-Lagrange equation

Consider the following functional (function of functions)

$$\boxed{F\left(f', f, x\right) = \int \phi\left(f'(x), f(x), x\right) dx} \tag{2}$$

Requiring that functional derivative of $F$ with respect to $f$ is zero at extrema (note $f$ is a function not a variable, hence the name functional derivative) following Eauler-Lagrange equation can be derived.

$$\boxed{\frac{\partial \phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial \phi}{\partial y'}\right) = 0} \tag{3}$$

Derivation of the above equation that is easy to follow is here

## Finding Stationary Point of Lagrangian in Eq 1.

Using the above Euler-Lagrange equation, the stationary conditions for the Lagrangian in equation 1, can be derived from functional derivative with respective $p$ and the derivative with respective $\lambda$ as follows.

$$\boxed{f(x) + T\left\{1 + ln(p(x))\right\} + \lambda = 0} \tag{4}$$

Setting derivative with respect to $\lambda$ to zero yields the contraint that PDF integral should be 1.

$$\boxed{\int p(x)\, dx - 1 = 0} \tag{5}$$

Rearranging eqn.4

$$\boxed{p(x) = e^{-(1+\lambda/T)}\, e^{-f(x)/T}} \tag{6}$$

Combining eqns. 5 and 6 the value of lambda can be expressed as follows

$$\boxed{\lambda = T\left\{ln \int e^{-f(x)/T} dx - 1\right\}} \tag{7}$$

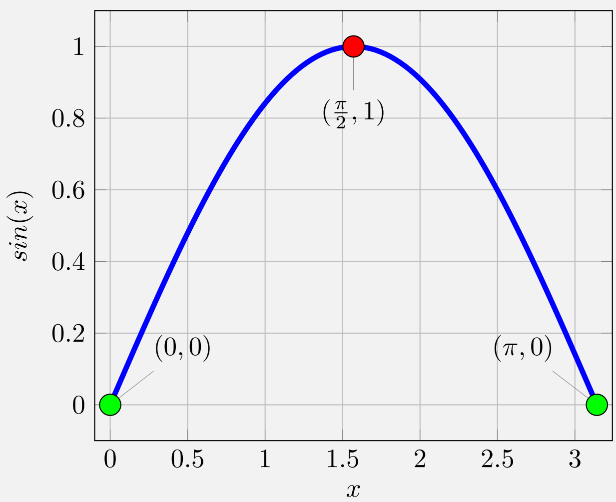
The question to ask is how do we now proceed to get min or max of $f(x)$ given that we have it in analytical form (eg. sin(x)).. What will be the p(x) as T approaches zero. That is the intent of the annealing process as we converge on p(x).

## Explorations with sinx

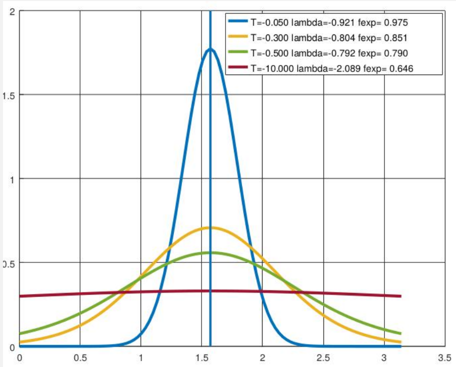
Assuming that we are interested in finding extrema of sin(x) in the interval $[0, \pi]$, one can proceed as follows.

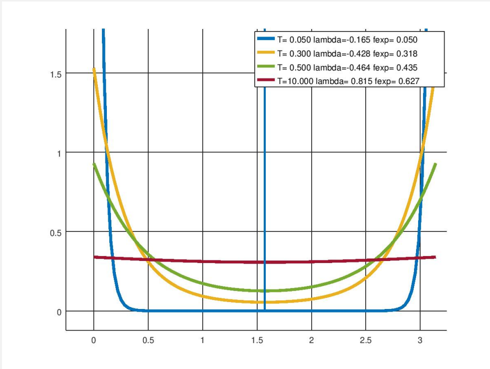
1. For a given T, calculate $\lambda$ by evaluating the integral numerically
2. With $Tand\lambda$ known, the PDF $p(x)$ now is well defined (eq.6). The expected value of $f(x)$ can be determined by numerical integration of $\int p(x)f(x)dx$

## $sin(x)$



## Change in pdf with T when $T \leq 0$ and $T \geq 0$




From the above figures, it is clear that for $T \geq 0$ as its magnitude approaches zero, the pdf spread decreases and tends to peak around the true maxima. On the otherhand, when $T \geq 0$, the pdf peaks appear near the end points where the function has minimal value in the interval of interest.

## $y(x) = 8(1 - 2x^2)x^2$