ANALYSIS AND OPTIMIZATION: CALCULUS OF VARIATIONS

SPRING 2016

In the optimization problems we have discussed so far, our goal was to find numerical values of a number of variables $x_1, \dots, x_n$ that optimize a certain quantity F that depends on $x_1, \dots, x_n$. We now discuss optimization problems whose goal is to find *functions* $x_1, \dots, x_n$ that opitimize a certain quantity F that depends on $x_1, \dots, x_n$.

1. Examples

To convince you that these problems arise naturally, here are some examples.

1.1. Production planning.

A firm receives an order to produce B units of a good in time T. They would like to fulfill the order in the lowest cost. The cost comes from two sources

(1) Storage costs amount a per unit good per unit time.
(2) Production at rate r costs amount b per unit good.

How should the firm plan its production so that the total cost is minimized?

Let us translate the problem into a purely mathematical problem. Denote by $x(t)$ the amount produced up to time t. This is the function that we want to find. Let us write the total cost incurred in terms of $x(t)$. For a small time period $[t, t+\Delta t]$, the cost incurred is approximately the following:

(1) Storage: $a \cdot x(t) \cdot \Delta t$
(2) Production: $b \cdot x'(t) \cdot x'(t)\Delta t$.

Summing over the entire interval and taking the limit as $\Delta t \to 0$, we arrive at the total cost

$$\text{Total cost} = \int_0^T ax(t) + bx'(t)^2 \, dt.$$

The optimization problem is then the following:

Problem 1. *Find a function $x(t)$ that minimizes*

$$\int_0^T ax(t) + bx'(t)^2 \, dt,$$

subject to

$$x(0) = 0 \quad x(T) = B.$$

1.2. The brachistochrone problem.

This problem in physics spawned the development of calculus of variations. The goal is to design a path between two given points A and B such that a ball that slides along this path under gravity takes the least time to go from A to B (see Figure 1).

Let us translate the problem into a purely mathematical problem. Suppose the coordinates of A are $(0, a)$ and the coordinates of B are $(b, 0)$. Suppose the path is given by the graph of a function y of x. Let us write the time taken for the ball to roll along this path. Suppose the ball is at (x,y) and rolls to $(x+\Delta x, y-\Delta y)$ in time Δt. Assume all the Δ's are small. The distance travelled is

$$\sqrt{\Delta x^2+\Delta y^2}=\sqrt{1+(y')^2}\cdot\Delta x.$$

The speed v is given by equating the kinetic energy with the loss in the potential energy (we are assuming that the initial speed is zero):

$$\frac{1}{2}mv^2 = mg(a-y)$$
$$v=\sqrt{2g(a-y)}.$$

So the time taken is

$$\Delta = \frac{\sqrt{1+(y')^2}\cdot\Delta x}{\sqrt{2g(a-y)}}.$$

Summing over the whole interval and taking the limit as $\Delta x\to 0$ gives the total time

$$\text{Total time} = \int_0^b \sqrt{\frac{1+(y')^2}{2g(a-y)}}\,dx.$$

This leads to the following problem.

Problem 2. *Find a function $y(x)$ that minimizes*

$$\int_0^b \sqrt{\frac{1+(y')^2}{2g(a-y)}}\,dx$$

subject to

$$y(0)=a \quad y(b)=0.$$

Exercise 3. How would the problem change if the ball was launched with some non-zero initial speed v_0?

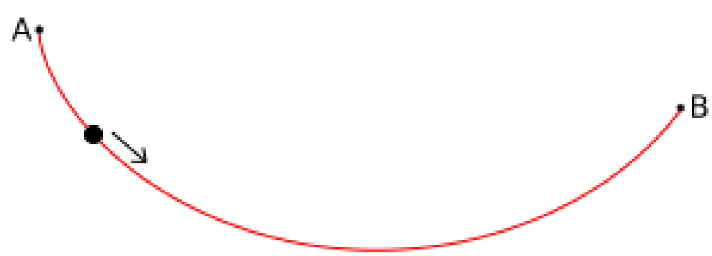

FIGURE 1. Which path leads to the quickest slide?

1.3. **Investment strategy.** Suppose we are planning an investment and consumption strategy, in the following setup. A capital K yields returns at a rate $F(K)$. The returns can be spent or re-invested, or a combination of both. Spending results in enjoyment, or utility, and re-investment results in greater K for greater future returns. Suppose consumption at rate C yields utility $U(C)$. Our goal is to devise a consumption/re-investment strategy that will maximize total utility.

Denote by $K(t)$ the invested capital at time t, by $C(t)$ the rate of consumption at time t and by $R(t)$ the rate of re-investment at time t. Then we have the equation

$$F(K(t)) = C(t) + R(t),$$

which expresses the fact that the returns from capital $K(t)$ are partially consumed and partially re-invested. Furthermore, the rate of re-investment is precisely the rate of growth of $K(t)$. Therefore, we have

$$K'(t) = R(t).$$

Using the two equations above, we can write

$$C(t) = F(K(t)) - K'(t).$$

We are given the initial value $K(0)$ and the utility function U. Our goal is to figure out the function $K(t)$ that leads to the maximum total utility

$$\int_0^t U(C(t))\, dt.$$

Therefore the problem is the following.

Problem 4. *Suppose K_0, U, F, and T are given. Find a function $K(t)$ that maximizes*

$$\int_0^T U\left(F(K(t)) - K'(t)\right)\, dt,$$

subject to $K(0) = K_0$.

Here is a specific example of the above. Suppose the rate of returns $F(K)$ is linear, say $F(K) = aK$, and the utility function is $U(C) = \ln C$. (Typically, the utility function will be increasing but concave, reflecting the diminishing marginal utility of consumption.) Then we want to find the $K(t)$ that maximizes

$$\int_0^T \ln(aK(t) - K'(t))\, dt.$$

To get some feel for the above, suppose we decide to re-invest the entire returns. Then we have $C(t) = 0$ and therefore $K'(t) = cK(t)$. The function that satisfies this equation, subject to the initial condition that $K(0) = K_0$ is $K(t) = K_0 e^{ct}$. So, if we re-invest the entire returns then the capital grows exponentially but the overall utility is zero. This is clearly not the optimum.

Suppose we decide to spend the entire returns. Then we have $C(t) = cK(t)$ and $K'(t) = 0$. Therefore, $K(t)$ is constant, namely K_0. The overall utility is

$$\int_0^T \ln(cK_0)\, dt = \ln(cK_0) \cdot T,$$

which is nonzero but may not be optimal.

2. The general problem

Suppose we are given real numbers a, b, A, B, and a function $F(u,v,w)$. Set

$$F_1(u,v,w) = \frac{\partial F}{\partial u}(u,v,w),$$
$$F_2(u,v,w) = \frac{\partial F}{\partial v}(u,v,w),$$
$$F_3(u,v,w) = \frac{\partial F}{\partial w}(u,v,w).$$

Let us denote by $\dot{x}$ the derivative $\frac{dx}{dt}$, by $\ddot{x}$ the double derivative $\frac{d^2x}{dt^2}$, and so on. Assume that all functions under considerations behave nicely (for example, are infinitely differentiable).

Problem 5 (General variations problem). *Find a function $x(t)$ that maximizes or minimizes*

$$\int_a^b F(t,x,\dot{x})\,dt$$

subject to $x(a) = A$ and $x(b) = B$.

3. The Euler–Lagrange equation

Theorem 6. *If x achieves the optimum in Problem 5, then it satisfies the Euler–Lagrange equation*

$$F_2(t,x,\dot{x}) - \frac{d}{dt}\big(F_3(t,x,\dot{x})\big) = 0.$$

Proof. Let h be any function with $h(a) = h(b) = 0$. Then for any ϵ, the function $x + \epsilon h$ also satisfies the boundary conditions. Let

$$G(\epsilon) = \int_a^b F(t, x+\epsilon h, \dot{x}+\epsilon\dot{h})\,dt.$$

Since x is an opitimum solution to Problem 5, $\epsilon = 0$ is the optimum value of G. Therefore, we must have $\frac{dG}{d\epsilon}(0) = 0$. Let us calculate this derivative:

$$\begin{aligned}
\frac{dG}{d\epsilon} &= \frac{d}{d\epsilon}\int_a^b F(t, x + \epsilon h, \dot{x} + \epsilon \dot{h})\, dt \\
&= \int_a^b \frac{d}{d\epsilon} F(t, x + \epsilon h, \dot{x} + \epsilon \dot{h})\, dt \\
&= \int_a^b F_1(t, x, \dot{x}) \cdot 0 + F_2(t, x, \dot{x})h + F_3(t, x, \dot{x})\dot{h}\, dt \quad \text{at } \epsilon = 0 \text{ by the chain rule} \\
&= \int_a^b F_2(t, x, \dot{x})h\, dt + F_3(t, x, \dot{x})h(t)|_a^b - \int_a^b \frac{d}{dt}\left(F_3(t, x, \dot{x})\right) h\, dt \quad \text{by parts} \\
&= \int_a^b \left(F_2(t, x, \dot{x}) - \frac{d}{dt}\left(F_3(t, x, \dot{x})\right)\right) h\, dt \quad \text{since } h(a) = h(b) = 0. \\
&= 0.
\end{aligned}$$

Now, it is easy to see that if $\int_a^b f(t)h(t)\, dt = 0$ for *all* functions $h(t)$ satisfying $h(a) = h(b) = 0$, then we must have $f(t) = 0$. Therefore, we conclude that

$$F_2(t, x, \dot{x}) - \frac{d}{dt}\left(F_3(t, x, \dot{x})\right) = 0.$$

□

4. Using the Euler–Lagrange equation

Let us use the Euler–Lagrange equation to solve (some of) the problems posed in § 1.

4.1. **Production planning.**

Problem 7. *Find a function $x(t)$ that minimizes*

$$\int_0^T ax + b\dot{x}^2\, dt,$$

subject to

$$x(0) = 0 \quad x(T) = B.$$

Solution. We have $F(u, v, w) = av + bw^2$, so $F_2(u, v, w) = a$ and $F_3(u, v, w) = 2bw$. The Euler–Lagrange equation is

$$\begin{aligned}
0 &= F_2(t, x, \dot{x}) - \frac{d}{dt} F_3(t, x, \dot{x}) \\
&= a - \frac{d}{dt}(2b\dot{x}) \\
&= a - 2b\ddot{x}.
\end{aligned}$$

That is, $\ddot{x} = a/2b$. By integrating twice, we get

$$x(t) = (a/4b)t^2 + ct + d,$$

where c and d are constants to be determined using the boundary conditions. The condition $x(0)=0$ gives $d=0$. The condition $x(T)=B$ gives

$$c=\frac{B}{T}-\frac{aT}{4b}.$$

So we get

$$x(t)=\frac{a}{4b}t^2+\left(\frac{B}{T}-\frac{aT}{4b}\right)t.$$

□

Remark 8. Strictly speaking, the only conclusion we can draw from solving the Euler–Lagrange equation is that *if* a minimizer $x(t)$ exists, then it must be the one we found. But let us assume that it does (which we will not prove).

Note some qualitative features of the solution. If $a=0$ (no storage cost), then the optimal strategy is to produce at a constant rate, which means that the graph of $x(t)$ is a straight line. If $a>0$, then the graph of $x(t)$ is a parabola, which gets steeper as $a/4b$ increases. So, for $a \gg b$ (negligible production cost compared to storage), the optimal strategy is to do most of the production close to the deadline.

4.2. **Investment planning.** Let us take the rate of returns to be proportional to the capital and the utility function to be the logarithm. Let us also assume that both the initial and the final value of the capital is given. For simplicity, let us assume $T=1$.

Problem 9. *Find a function $K(t)$ that maximizes*

$$\int_0^1 \ln(cK-\dot{K})\,dt,$$

subject to $K(0)=K_0$ and $K(1)=K_1$.

Solution. Here $F(u,v,w)=\ln(cv-w)$, So

$$F_2(u,v,w)=c/(cv-w)$$
$$F_3(u,v,w)=-1/(cv-w).$$

The Euler–Lagrange equation becomes

$$\begin{aligned}0&=\frac{c}{cK-\dot{K}}-\frac{d}{dt}\left(\frac{-1}{cK-\dot{K}}\right)\\&=\frac{c}{cK-\dot{K}}-\frac{c\dot{K}-\ddot{K}}{(cK-\dot{K})^2}\\0&=c(cK-\dot{K})-c\dot{K}+\ddot{K}\\0&=\ddot{K}-2c\dot{K}+c^2K.\end{aligned}$$

The solution to this differential equation is

$$K(t)=Ae^{ct}+Bte^{ct},$$

for constants A and B to be determined from the boundary conditions. The consumption function is

$$C(t)=-Be^{ct}.$$

Let us look at two interesting special cases. The first is $K_1 = 0$. This will be the case if there are no financial obligations at $t = 1$. In this case $A = K_0$ and $A + B = 0$, so $B = -K_0$. Therefore, we have

$$K(t) = K_0(1 - t)e^{ct}.$$

The consumption function is

$$C(t) = K_0 e^{ct}.$$

The second case is $K_1 = K_0$. This will be the case if the initial capital is to be returned at $t = 1$. In this case $A = K_0$ and $e^c(A + B) = K_0$, so $B = -(1 - e^{-c})K_0$. Therefore, we have

$$K(t) = K_0 e^{ct} \left(1 - (1 - e^{-c})t\right).$$

The consumption function is

$$C(t) = (1 - e^{-c})K_0 e^{ct}.$$

Interestingly, the consumption function is exponential in both cases. In the second case, only the leading coefficient is slightly smaller.

Let us compare the optimum strategy in the second case with the naïve strategy of consuming all the returns immediately. In both cases, we have $K(0) = K(1) = K_0$. The total utility in the naïve strategy is

$$\ln(cK_0) = \ln c + \ln K_0.$$

The total utility in the optimum strategy is

$$\int_0^1 \ln(1 - e^{-c}) + \ln K_0 + ct \; dt = \ln(1 - e^{-c}) + \ln K_0 + c/2.$$

Can you see that the optimum strategy outperforms the naïve one? □