

Optimal Control for Dummies

Marcus Davidsson
Newcastle University, UK

Abstract: The aim of this study is to facilitate the teaching of optimal control in economics. Understanding optimal control can be challenging to say the least. This study discusses optimal control and its applications in economics. More specifically the famous Keynesian-Ramsey rule will be explored and discussed in detail. The same rule is also extensively used in resource economics to calculate the optimal extraction from for example a gold mine.

Keywords: Dynamic optimization, growth, macro economics, optimal control

INTRODUCTION

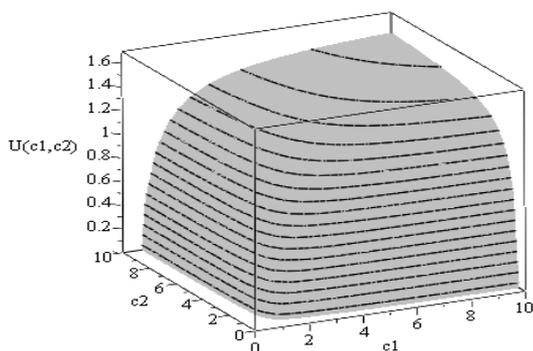
Optimal control theory was introduced to the world by the Russian mathematician Lev Pontryagin in 1962. Lawrence (2002) explains that Lev went blind at the age of fourth teen which makes the achievement even more remarkable. Optimal control is used to solve dynamic optimization problems that frequently arise in engineering and economic applications. Optimal control problems can also be expressed as constrained optimization problems (Lagrange) or as dynamic programming problems. Dynamic programming was introduced by Bellman (1954) and represents as well an important component of dynamic optimization. Optimal control can be applied to many different problems. Yang *et al.* (1987) apply optimal control to model earthquakes. The authors explain that the Riccati closed-loop control does not satisfy the optimality condition due to continuous external disturbances such as ground acceleration. They therefore propose a new optimal control routine which includes an instantaneous optimal closed-loop control in order to solve such a discrepancy. Hinze and Kunisch (2001) look at how optimal control is applied in fluid dynamics. They use numerical examples to show that optimal control can be used to solve the classic Navier-Stokes equations by second order numerical methods. They use both Newton and quasi-Newton methods to solve the optimal flow control problem. Sohngen and Mendelsohn (2002) apply optimal control to greenhouse gas emission. They show that the rental price for carbon sequestration will rise over time given that carbon accumulates in the atmosphere. Landowners can decrease the amount of carbon by increasing forestland and lengthening rotations. Tropical forests store over two-thirds of this added carbon. Optimal control has also extensively been applied to economics (Barro and Sala-i-Martin, 2003).

BASIC THEORETICAL MODELS

In this section we will discuss dynamic optimization and its applications in economics. The classic example in economics is the capital-consumption problems i.e., how much should we consume in order to maximize our utility from consumption given a fixed budget constraint. The same approach is also extensively used in resource economics to calculate the optimal extraction from for example a gold mine. We will use 4 different approaches to dynamic optimization as seen below. All of these different approaches will result in the famous Keynesian-Ramsey rule (Ramsey, 1928).

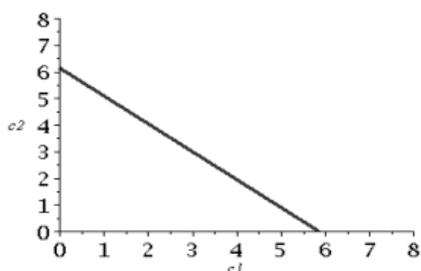
- Constrained Optimization in the form of Lagrange (discrete time)
- Calculus of Variation (continuous time)
- Optimal Control; Present Value Hamiltonian (continuous time)
- Optimal Control; Current Value Hamiltonian (continuous time)

In this initial example let assume that we have a two time period economy and that we have an individual that gets utility from consumption $u(c_1, c_2)$ where c_1 is the amount of consumption in period one and c_2 is the amount of consumption in period 2. We also assume that the individual prefers consumption today over consumption in the future hence he will discount utility from consumption over time where the discount factor is given by $B = 1/(1 + r)$ The utility function as presented below is a so called Constant Intertemporal Elasticity of Substitution (CIES) utility function where the elasticity parameter θ determines the shape of the utility function i.e., how much consumption in period 1 is worth compared to consumption in period 2. Such a utility function is plotted in Fig. 1.



$$u(c1, c2) = \frac{c1^{1-\theta} - 1}{1-\theta} + B * \frac{c2^{1-\theta} - 1}{1-\theta}$$

Fig. 1: Two time period CIES utility function



$$c2 = k1 * (1 + r) + k2 - (1 + r) * c1$$

Fig. 2: Budget constraint

Since this individual does not have an infinite amount of money he also has a budget constraint given by $k1 + k2 = c1 + c2$ which simply says that the total amount capital in the two time periods must be equal to the total amount of consumption in the two time periods. We again assume that this individual prefers capital and consumption today over capital and consumption in the future hence he will discount capital and consumption over time which means that his

budget constraint is given by $k1 + k2 / (1 + r) = c1 + c2 / (1 + r)$. If we solve for $c2$ and multiply both sides by $(1+r)$ such an equation can be expressed in Fig. 2. Note that $(1+r)$ is now the slope the budget constraint.

The question now becomes how much should such an individual consume in the two time periods in order to maximize his utility from consumption? We can find the answer to such a question by setting up the Lagrange (constrained optimization) equation where the objective is to maximize utility from consumption given the budget constraint. Note also that the budget constraint is expressed in an equal to zero form as:

$$L = \frac{c1^{1-\theta}-1}{1-\theta} + B * \frac{c2^{1-\theta}-1}{1-\theta} + \lambda * \left(k1 + \frac{k2}{(1+r)} - c1 - \frac{c2}{(1+r)} \right)$$

$$\frac{\partial L}{\partial c1} = 0 \rightarrow \frac{c1^{1-\theta}}{c1} - \lambda = 0 \rightarrow$$

$$i) \lambda = \frac{c1^{1-\theta}}{c1}$$

$$\frac{\partial L}{\partial c2} = 0 \rightarrow \frac{Bc2^{1-\theta}}{c2} - \frac{\lambda}{(1+r)} = 0 \rightarrow$$

$$ii) \lambda = \frac{Bc2^{1-\theta}(1+r)}{c2}$$

Since both i) and ii) are equal to λ we can set them equal to each other which gives us:

$$\frac{c1^{1-\theta}}{c1} = \frac{Bc2^{1-\theta}(1+r)}{c2} \rightarrow \frac{c1^{1-\theta}}{Bc2^{-\theta}} = (1+r)$$

We know since previously that $(1+r)$ is simply the slope of the budget constraint. It turns out that the expression on the left -hand is simply the slope the utility function. At the optimal point these two equations must be equal to each other as seen in Fig. 3.

The optimality condition can be written as:

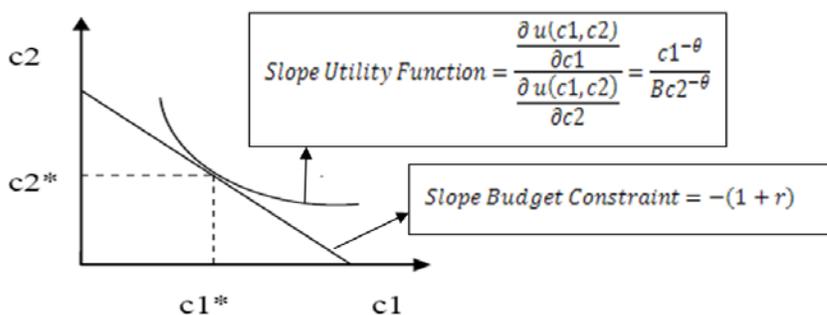


Fig. 3: Consumption in equilibrium

$$\begin{aligned} \frac{c1^{-\theta}}{Bc2^{-\theta}} &= (1+r) \rightarrow \left(\frac{c1}{c2}\right)^{-\theta} = B(1+r) \rightarrow \left(\frac{c1}{c2}\right)^{-\theta} \\ &= \frac{(1+r)}{(1+p)} \rightarrow \left(\frac{c2}{c1}\right)^{\theta} = \frac{(1+r)}{(1+p)} \rightarrow \frac{c2}{c1} \\ &= \left(\frac{(1+r)}{(1+p)}\right)^{\frac{1}{\theta}} \rightarrow \ln\left(\frac{c2}{c1}\right) = \ln\left(\left(\frac{(1+r)}{(1+p)}\right)^{\frac{1}{\theta}}\right) \\ &\rightarrow \ln\left(\frac{c2}{c1}\right) = \frac{1}{\theta} \ln\left(\frac{(1+r)}{(1+p)}\right) \\ &\rightarrow \ln(c2) - \ln(c1) = \frac{\ln(1+r) - \ln(1+p)}{\theta} \end{aligned}$$

We should now note two things:

- i) $\ln(1+z) \approx z$ for example $\ln(1+0.05) = 0.04879$
- ii) $c2 = c1 * \exp(g) \rightarrow \left(\frac{c2}{c1}\right) = \exp(g) \rightarrow \ln\left(\frac{c2}{c1}\right) = \ln(\exp(g)) \rightarrow \ln(c2) - \ln(c1) = g$

where, g is the percentage growth when consumption is growing exponentially. The percentage growth rate is defined as:

$$g = \lim_{\Delta t \rightarrow 0} \left(\frac{c(t+\Delta t) - c(t)}{c(t)} \right) = \frac{\dot{c}(t)}{c(t)}$$

This means that our previous equation can be written as:

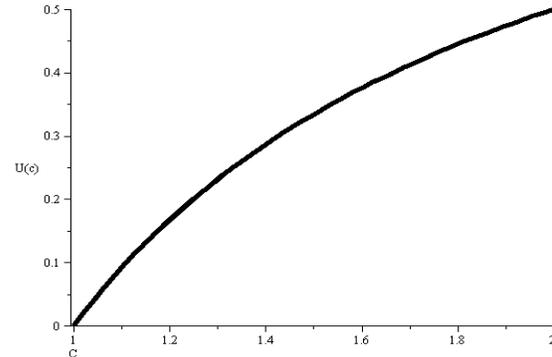
$$\frac{\dot{c}(t)}{c(t)} = \frac{r-p}{\theta}$$

This equation is called the Keynesian-Ramsey rule. Such equation of motion of consumption tells us how much consumption will change over time given that that the consumer maximizes his utility from consumption subject to the capital constraint.

In calculus of variation the previous problem looks a little bit different. The end result will be the same though. The objective is to maximize discounted utility from consumption $\exp(-p * t) * u(c)$ given the budget constraint $\dot{k}(t) = w * L + r * k - c$ which says that the rate of change in capital is given by how much you work times the wage rate plus the return on capital minus consumption. Again the constraint is expressed in an equal-to-zero form. The optimization problem has the following form:

$$\begin{aligned} L &= \exp(-p * t) * u(c(t)) + \lambda(t) \\ &* (w * L + r * k(t) - c(t) - \dot{k}(t)) \end{aligned}$$

Now instead of deriving the Euler equation by using integration by parts (Appendix) we will reverse engineer the process and simply argue; what type of equalities must be satisfied in order for our problem to



$$u(c(t)) = \frac{c(t)^{1-\theta} - 1}{1-\theta}$$

Fig. 4: One period CIES utility function

produce the Keynesian-Ramsey equation? Note that the Euler equation is the link between Lagrange (discrete time) and Optimal Control (continuous time). The Euler equation (calculus of variation) specifies the property of an extremal. Without this equation we cannot explain the structure of the current and present value Hamiltonian. It turns out that the following relationship must hold:

$$\frac{\partial L}{\partial k} = \frac{d}{dt} \left[\frac{\partial L}{\partial \dot{k}(t)} \right] \rightarrow \frac{\partial L}{\partial k} = -\frac{\lambda(t)}{dt} = -\dot{\lambda}(t)$$

This means that all first order conditions are not set to zero as previously:

$$\begin{aligned} \frac{\partial L}{\partial k} &= -\dot{\lambda}(t) \rightarrow \lambda(t) * r = -\dot{\lambda}(t) \rightarrow \\ i) \quad \frac{\dot{\lambda}(t)}{\lambda(t)} &= -r \\ \frac{\partial L}{\partial c} &= 0 \rightarrow \exp(-p * t) * \frac{\partial u(c(t))}{\partial c(t)} - \lambda = 0 \\ \rightarrow ii) \quad \lambda(t) &= \exp(-p * t) * \frac{\partial u(c(t))}{\partial c(t)} \end{aligned}$$

We now note that the equation for the CIES utility function is given in Fig. 4.

This means that the partial derivative of the utility function with respect to consumption is given by:

$$\frac{\partial u(c(t))}{\partial c(t)} = c(t)^{-\theta}$$

We can plug in such an expression in equation ii) and take the logarithm on both sides:

$$\begin{aligned} \lambda(t) &= \exp(-p * t) * c(t)^{-\theta} \rightarrow \ln(\lambda(t)) \\ &= \ln(\exp(-p * t) * c(t)^{-\theta}) \rightarrow \ln(\lambda(t)) \\ &= -p * t - \theta * \ln(c(t)) \end{aligned}$$

We differentiate with respect to time which gives us:

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -p - \theta * \frac{\dot{c}(t)}{c(t)}$$

We now set such an equation equal to i) which gives us:

$$-r = -p - \theta * \frac{\dot{c}(t)}{c(t)} \rightarrow \frac{\dot{c}(t)}{c(t)} = \frac{r-p}{\theta}$$

which again is the Keynesian-Ramsey rule i.e., the same equation that we had in the Lagrange example. Such an equation was found by simply setting the partial derivative with respect to capital equal to $-\dot{\lambda}(t)$ and the partial derivative with respect to consumption equal to zero. Now we know that the Lagrange and calculus of variation examples produces the same result.

We can now look that optimal control by introducing the Present Value Hamiltonian. The optimization problem has the following form:

$$H = \exp(-p * t) * u(c(t)) + \lambda(t) * (w * L + r * k(t) - c(t))$$

Now the partial derivatives take the same form as in the calculus of variation example

$$\frac{\partial H}{\partial k} = -\dot{\lambda}(t) \rightarrow \lambda(t) * r = -\dot{\lambda}(t) \rightarrow$$

$$i) \frac{\dot{\lambda}(t)}{\lambda(t)} = -r$$

$$\frac{\partial H}{\partial c} = 0 \rightarrow \exp(-p * t) * \frac{\partial u(c(t))}{\partial c(t)} - \lambda = 0 \rightarrow$$

$$ii) \lambda(t) = \exp(-p * t) * \frac{\partial u(c(t))}{\partial c(t)}$$

Again we plug in the expression for the partial derivative of the utility function with respect to consumption in equation ii) and take the logarithm on both sides:

$$\begin{aligned} \lambda(t) &= \exp(-p * t) * c(t)^{-\theta} \rightarrow \ln(\lambda(t)) \\ &= \ln(\exp(-p * t) * c(t)^{-\theta}) \rightarrow \ln(\lambda(t)) \\ &= -p * t - \theta * \ln(c(t)) \end{aligned}$$

We differentiate with respect to time which gives us:

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -p - \theta * \frac{\dot{c}(t)}{c(t)}$$

We now set such an equation equal to i) which gives us:

$$-r = -p - \theta * \frac{\dot{c}(t)}{c(t)} \rightarrow \frac{\dot{c}(t)}{c(t)} = \frac{r-p}{\theta}$$

which again is the Keynesian-Ramsey rule.

Finally we will look at the current value Hamiltonian. The optimization problem has the following form:

$$H = u(c(t)) + \lambda(t) * (w * L + r * k(t) - c(t))$$

We should first note that the discounting term $\exp(-p * t)$ is missing in the Hamiltonian. This means that we need to do the discounting in the first order conditions as seen below. Note that when $p = 0$ then the first order condition is simply reduced to the first order conditions for the present value Hamiltonian $-\dot{\lambda}(t)$.

$$\frac{\partial H}{\partial k} = p * \lambda(t) - \dot{\lambda}(t) \rightarrow \lambda(t) * r$$

$$= p * \lambda(t) - \dot{\lambda}(t) \rightarrow i) -\frac{\dot{\lambda}(t)}{\lambda(t)} = r - p$$

$$\frac{\partial H}{\partial c} = 0 \rightarrow \frac{\partial u(c(t))}{\partial c(t)} - \lambda = 0 \rightarrow$$

$$ii) \lambda(t) = \frac{\partial u(c(t))}{\partial c(t)}$$

Again we plug in the expression for the partial derivative of the utility function with respect to consumption in equation ii) and take the logarithm on both sides:

$$\begin{aligned} \lambda(t) &= c(t)^{-\theta} \rightarrow \ln(\lambda(t)) = \ln(c(t)^{-\theta}) \\ &\rightarrow \ln(\lambda(t)) = -\theta * \ln(c(t)) \end{aligned}$$

We differentiate with respect to time which gives us:

$$-\frac{\dot{\lambda}(t)}{\lambda(t)} = \theta * \frac{\dot{c}(t)}{c(t)}$$

We now set such an equation equal to i) which gives us:

$$r - p = \theta * \frac{\dot{c}(t)}{c(t)} \rightarrow \frac{\dot{c}(t)}{c(t)} = \frac{r-p}{\theta}$$

which again is the Keynesian-Ramsey rule.

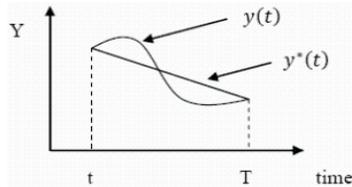
DISCUSSION AND CONCLUSION

Dynamic optimization or more specifically optimal control has been studied in detail. Optimal control requires a fair amount of mathematics and equation manipulation which means that optimal control might not be as accessible to the average person as other more straight-forward economic topics. Despite such a fact optimal control represent a cornerstone in economic theory which means that it is important to understand the basic concept in order to create your own

mathematical models which is critical for economic research. Hopefully this study has managed to present optimal control in an understandable way to help facilitate such a process. Note that all mathematical derivations in this study have been done with the mathematical software Maple. Maple is a good tool to use to make sure you get all the differentiating correct but it can also be a challenging tool when it comes to equation manipulation. The focus of this study has been on presenting a backward inductive and intuitive way of deriving the Keynesian-Ramsey rule. Four different ways of deriving the Keynesian-Ramsey rule was presented. Such an intuitive approach was chosen due to the added complexity of for example integration by parts.

Appendix: Property of an extremal:

The Euler equation: An extremal is as a path that minimizes or maximizes our objective function. We define $y^*(t)$ = extremal and $\dot{y}^*(t)$ = **slop extremal** For example, we know that the shortest distance between two points is a straight line. This straight line would be defined as an extremal. The basic question we need to answer is therefore: What property does this straight line have that the neighboring curve $y(t)$ does not have? The answer to this question is the Euler equation, which is derived in this section (Chiang, 1992).



We now introduce a small deviation to our extremal in the form of $\epsilon \cdot p(t)$ and $\epsilon \cdot \dot{p}(t)$ This gives us non-extremal neighboring path and path slopes defined as:

$$y(t) = y^*(t) + \epsilon \cdot p(t)$$

and

$$\dot{y}(t) = \dot{y}^*(t) + \epsilon \cdot \dot{p}(t)$$

Note that $\dot{y}(t)$ is the slope of $y(t)$ which also can be written as $\text{diff}(y(t), t)$ Note also that when $\epsilon \rightarrow 0$ then $y(t) \rightarrow y^*(t)$ and $\dot{y}(t) \rightarrow \dot{y}^*(t)$ This means that when $\epsilon \rightarrow 0$ the neighboring path becomes the extremal. This means that we can displace the extremal to neighboring positions by changing the magnitude of ϵ . Note also that $y(t) \rightarrow y^*(t)$ at the end points which means that $\epsilon = 0$ which means that $p(t)$ must also be equal to zero at these points. We now have all the necessary information so we can set up the functional form of calculus of variation as follows: Our objective is to find the optimal path of $y(t)$ that maximize the following integral:

$$V(\epsilon) = \int_t^T F(y(t), \dot{y}(t)) dt$$

We now take the derivative of $V(\epsilon)$ with respect to ϵ by using the multivariable chain rule:

$$\frac{dF(x(t), y(t))}{dt} = \frac{\partial F(\dots)}{\partial x(t)} \frac{dx(t)}{dt} + \frac{\partial F(\dots)}{\partial y(t)} \frac{dy(t)}{dt}$$

This gives us:

$$\frac{dV(\epsilon)}{d\epsilon} = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} \frac{dy(t)}{d\epsilon} + \frac{\partial F(\dots)}{\partial \dot{y}(t)} \frac{d\dot{y}(t)}{d\epsilon} dt$$

We now note that $y(t) = y^*(t) + \epsilon \cdot p(t)$ which means that:

$$\text{diff}(y^*(t) + \epsilon \cdot p(t), \epsilon) p(t)$$

We also note that $\dot{y}(t, \epsilon) = \dot{y}^*(t) + \epsilon \cdot \dot{p}(t)$

which means that:

$$\text{diff}(\dot{y}^*(t) + \epsilon \cdot \dot{p}(t), \epsilon)$$

$$\frac{d}{dt} p(t)$$

which can be written as $\dot{p}(t)$ If we plug these two expression into our pervious equation we get:

$$\frac{dV(\epsilon)}{d\epsilon} = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} p(t) + \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt$$

We now note that in order for us to have an extremal then ϵ must be equal to zero. This means that $\text{diff}(V(\epsilon), \epsilon)$ must also be equal to zero. This gives us:

$$0 = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} p(t) + \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt$$

which can be written as:

$$0 = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} p(t) dt + \int_t^T \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt$$

Note that we define the above equation as equation XX. We are now going to integrated the second term by parts:

$$\int_t^T \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt$$

We note that

#####

The product rule is given by:

$$\frac{\partial F(f(t)g(t))}{\partial t} = f(t) \frac{dg(t)}{dt} + g(t) \frac{df(t)}{dt}$$

We now integrate both sides:

$$\int_t^T \frac{\partial F(f(t), g(t))}{\partial t} dt = \int_t^T f(t) \frac{dg(t)}{dt} dt + g(t) \frac{df(t)}{dt} dt$$

Which can be written as:

$$\int_t^T \frac{\partial F(f(t), g(t))}{\partial t} dt = \int_t^T f(t) \frac{dg(t)}{dt} dt + \int_t^T g(t) \frac{df(t)}{dt} dt$$

We now note that if we have an integral of the form $\int_t^T F dt$ Then the derivative with respect to time of such an integral is given by the original function $\frac{d}{dt} \int_t^T F dt = F = F$ This means that if we evaluate the expression on the left hand side in the above equations at the end points t and T we can write the above equations as:

$$\left[F(f(t), g(t)) \right]_{t \text{ and } T} = \int_t^T f(t) \frac{dg(t)}{dt} dt + \int_t^T g(t) \frac{df(t)}{dt} dt$$

If we rearrange the above equation we can write it as:

$$\int_t^T f(t) \frac{dg(t)}{dt} dt = \left[F(f(t), g(t)) \right]_{t \text{ and } T} - \int_t^T g(t) \frac{df(t)}{dt} dt$$

We now define:

$$\frac{df(t)}{dt} dt = \dot{f}(t) dt = dV \quad \frac{dg(t)}{dt} dt = \dot{g}(t) dt = dU$$

Note that we start by defining the derivatives. This is important because it will help us derive the correct expression in our main example. This means that $\dot{f}(t) = V$ $\dot{g}(t) = U$ This gives us the expression for integration by parts:

$$\int_t^T V dU = [VU]_{t \text{ and } T} - \int_t^T U dV$$

###

As previously, we start by defining the derivative. We therefore get:

$$\frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} dt = dV \quad \frac{d [p(t)]}{dt} dt = \dot{p}(t) dt = dU$$

This means that,

$$\frac{\partial F(\dots)}{\partial \dot{y}(t)} = V \quad p(t) = U$$

This means that,

$$\int_t^T \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt = \left. \frac{\partial F(\dots)}{\partial \dot{y}(t)} p(t) \right|_{t \text{ and } T} - \int_t^T p(t) \frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} dt$$

We now note that that at the end points $p(t) = 0$. This means that the terms after the equality sign and before the minus sign disappear. This gives us:

$$\int_t^T \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt = - \int_t^T p(t) \frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} dt$$

If we now plug the above expression into equation XX given by:

$$0 = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} p(t) dt + \int_t^T \frac{\partial F(\dots)}{\partial \dot{y}(t)} \dot{p}(t) dt$$

we get:

$$0 = \int_t^T \frac{\partial F(\dots)}{\partial y(t)} p(t) dt - \int_t^T p(t) \frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} dt$$

Which can be written as:

$$0 = \int_t^T p(t) \left[\frac{\partial F(\dots)}{\partial y(t)} - \frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} \right] dt$$

In order for the above expression to be zero, it is sufficient if:

$$\left[\frac{\partial F(\dots)}{\partial y(t)} - \frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt} \right] = 0$$

This is our final expression, the Euler equation, the necessary condition for an extremal. The fancy expression for a straight line. We will use this equation later because the Euler equation is the link between Lagrange and Optimal Control. Without this equation we cannot explain the structure of the current and present value Hamiltonian. One final thing to note is that usually our integral has a much simpler functional form:

$$V(\epsilon) = \int_t^T F(y(t)) dt$$

In such a case $\dot{y}(t)$ is missing which means that $\dot{y}(t) = 0$ which means our expression:

$$\frac{d \left[\frac{\partial F(\dots)}{\partial \dot{y}(t)} \right]}{dt}$$

must also be equal to zero. This means that our Euler equation is simply reduced to:

$$\frac{\partial F(\dots)}{\partial y(t)} = 0$$

This reduced form Euler equation is simply equal to Lagrangian's first order conditions. This means that the Lagrange equation can more precisely be defined as the Euler-Lagrangian equation.

REFERENCES

- Barro, R. and X. Sala-i-Martin, 2003. *Economic Growth*. 2nd Edn., McGraw-Hill Inc., Singapore.
- Bellman, R., 1954. The theory of dynamic programming. *Bull. Am. Math. Soc.*, 60(6): 503-515.
- Chiang, A., 1992. *Elements of Dynamic Optimization*. McGraw Hill, New York.
- Hinze, M. and K. Kunisch, 2001. Second order methods for optimal control of time-dependent fluid flow. *Soc. Indus. Appl. Math. J. Contr. Optim.*, 40(3): 925-946.
- Lawrence, W., 2002. The World of Blind Mathematicians. *Notices Am. Math. Soc.*, 49(10): 1246-1251.
- Pontryagin, L., 1962. *The Mathematical Theory of Optimal Processes*. Gordon and Breach Science Publishers, New York.
- Ramsey, F., 1928. A Mathematical Theory of Saving. *Econ. J.*, 38(152): 543-559.
- Sohngen, B. and R. Mendelsohn, 2002. An optimal control model of forest carbon sequestration. *Am. J. Agric. Econ.*, 85(2): 448-457.
- Yang, J., A. Akbarpour and P. Ghaemmaghami, 1987. New optimal control algorithms for structural control. *J. Eng. Mech.*, 113(9).