Control optimization theory notes.

We will deal with the following optimization problem :

Consider a generic system $X[t]$ with n degrees of freedom

which evolves under the general equation

$$\dot{X}[t] = f[X[t], \alpha[t]] \equiv F$$

where f describes the evolution of the system, as a function

of a control $\alpha[t]$ onto which we can operate. For example,

$X[t]$ could describe a car moving on a road, and $\alpha[t]$ could

represent the position of the steering wheel.

Then, the state of the system depends in a predictable manner

on both the initial state of the system $X[0]$ and the future

control actions $\alpha[t]$ which will be applied onto it.

Our overall task will be to determine what is the “best” control for

our system, if we are interested in asking the resulting evolution

to maximize some property of the system ?

For this we need to specify a specific payoff (or reward) criterion.

Let us define the payoff functional

$$P[\alpha[.]] = \int_0^T r[x[t], \alpha[t]] \, dt + g[x[T]] \qquad (*)$$

Our aim is to find a control $\alpha^*[.]$, which maximizes the payoff.

Such a control $\alpha^*[.]$ is called optimal.

The main result of control theory is the theoretically interesting and practically useful theorem that if $\alpha^*[.]$ is an optimal control, then there exists a function $p^*[.]$ called the costate, that satisfies a certain maximization principle.

For the control $\alpha^*[t]$ which maximizes $P[\alpha[.]]$ (*)

$\boxed{\text{there exists a function } P[t] \text{ (**)}}$ which, together with the optimal evolution $X^*[t]$ produced by

$$\dot{X}[t] = f\left[X[t],\ \alpha^*[t]\right]$$

satisfy the set of equations

$$\dot{X} = \frac{\partial H}{\partial P}$$

$$\dot{P} = -\frac{\partial H}{\partial X}$$

where $H = H[X, P, a] = P \,.\, f + r \equiv$ control theory Hamiltonian of (*).

Finally $\alpha^*[t]$, the optimal control, takes the value a which maximizes $H(X, P, \alpha)$. (***)

Comments :

1. by construction, the first hamilton equation $\dot{X} = \frac{\partial H}{\partial P}$

is just a reformulation of $\dot{X}[t] = f[X[t], \alpha[t]] \equiv F$

2. The second hamilton equation $\dot{P} = -\frac{\partial H}{\partial X}$ enables us to

follow the evolution of the new auxiliary function P[t]

which then in turn provides $\alpha^*[t]$ (***).

Example :

We want to MINIMIZE

$$Q = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt$$

for a system evolving under

$\dot{x}[t] = x[t] + \alpha[t]$

$x[0] = x_0$

We will maximize $P = -\int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt$

with the initial condition $x[0] = 1$ and $T = 1$

We describe the system state as

$X[t] = x[t]$

The evolution is

$\dot{X} = x[t] + \alpha[t] \equiv F$

The co-state is $P[t] = p[t]$

(the number of Lagrange multipliers $p[t]$ being equal to the dimension of dynamics F).

The control hamiltonian is

$$H = p . F + r = p\,(x + \alpha) - \left(x^2 + \alpha^2\right)$$

The control $\alpha[t]$ which maximizes $P[\alpha[.]] = -\int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt$

is found via the set of equations

$\dot{X} = \dfrac{\partial H}{\partial P} \qquad \dot{P} = -\dfrac{\partial H}{\partial X}$ and $\alpha[t]$ takes

the value a which maximizes $H\,(x, p, \alpha) = p\,(x + \alpha) - \left(x^2 + \alpha^2\right)$

(where $p = p[t]$, $x = x[t]$) ie, $\alpha[t] = p[t]/2$

The evolution of the state and co-state $x[t]$ and $p[t]$ is then

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1/2 \\ 2 & -1 \end{pmatrix}}_{M} \begin{pmatrix} x \\ p \end{pmatrix}$$

with $x[0] = x_0$, $p[0] = \lambda$, unknown.

Formally,

$$\begin{pmatrix} x[t] \\ p[t] \end{pmatrix} = \mathrm{Exp}[M\,t].\begin{pmatrix} x[t] \\ p[0] \end{pmatrix} = \mathrm{Exp}[M\,t].\begin{pmatrix} x_0 \\ \lambda \end{pmatrix}$$

We find

$$\text{Exp}[M\,t] = \begin{pmatrix} \dfrac{e^{-\sqrt{2}\,t}\left(-1+\sqrt{2}+e^{2\sqrt{2}\,t}+\sqrt{2}\,e^{2\sqrt{2}\,t}\right)}{2\sqrt{2}} & \dfrac{\left(-1+\sqrt{2}\right)\left(1+\sqrt{2}\right)e^{-\sqrt{2}\,t}\left(-1+e^{2\sqrt{2}\,t}\right)}{4\sqrt{2}} \\ \dfrac{e^{-\sqrt{2}\,t}\left(-1+e^{2\sqrt{2}\,t}\right)}{\sqrt{2}} & \dfrac{e^{-\sqrt{2}\,t}\left(1+\sqrt{2}-e^{2\sqrt{2}\,t}+\sqrt{2}\,e^{2\sqrt{2}\,t}\right)}{2\sqrt{2}} \end{pmatrix}$$

and

$$\begin{pmatrix} x[t] \\ p[t] \end{pmatrix} = \begin{pmatrix} x0\,\text{Cosh}\left[\sqrt{2}\,t\right] + \dfrac{(2\,x0+\lambda)\,\text{Sinh}\left[\sqrt{2}\,t\right]}{2\sqrt{2}} \\ \lambda\,\text{Cosh}\left[\sqrt{2}\,t\right] + \dfrac{(2\,x0-\lambda)\,\text{Sinh}\left[\sqrt{2}\,t\right]}{\sqrt{2}} \end{pmatrix}$$

The optimal control therefore is

$$\alpha[t] = p[t]/2 = \frac{\lambda}{2}\,\text{Cosh}\left[\sqrt{2}\,t\right] + \frac{(x0-\lambda/2)\,\text{Sinh}\left[\sqrt{2}\,t\right]}{\sqrt{2}}$$

for some λ unknown and $x0 = 1$

$Q = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt$ can be considered a penalization function

which depends on λ. Let us compute the optimal value of both λ and Q.

Inserting $x[t] = \text{Cosh}\left[\sqrt{2}\,t\right] + \dfrac{(2+\lambda)\,\text{Sinh}\left[\sqrt{2}\,t\right]}{2\sqrt{2}}$ and

$$\alpha[t] = \frac{\lambda}{2}\,\text{Cosh}\left[\sqrt{2}\,t\right] + \frac{(x0-\lambda/2)\,\text{Sinh}\left[\sqrt{2}\,t\right]}{\sqrt{2}}$$

We can compute the quantity $Q = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) \, dt$ using the

above equations and we get

$$Q = \left(4 + 4\lambda - \lambda^2\right) \mathrm{Sinh}\left[\sqrt{2}\right] + \left(\frac{1}{\sqrt{2}} + \frac{\lambda^2}{4\sqrt{2}}\right) \mathrm{Cosh}\left[\sqrt{2}\right] \mathrm{Sinh}\left[\sqrt{2}\right]$$

The minimum corresponds to

$$\left\{\lambda \to -\frac{2\,\mathrm{Sinh}\left[\sqrt{2}\right]}{\sqrt{2}\,\mathrm{Cosh}\left[\sqrt{2}\right] - \mathrm{Sinh}\left[\sqrt{2}\right]} = -3.379\right\}$$

The solution found sets the minimum value for $Q = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) \, dt$

for a system following

$\dot{x}[t] = x[t] + \alpha[t]$

and initial conditions $x[0] = 1$

as $Q_{min} = 1.6895$

for the control $\alpha[t] = \dfrac{\mathrm{Sinh}\left[\sqrt{2}\,(-1+t)\right]}{\sqrt{2}\,\mathrm{Cosh}\left[\sqrt{2}\right] - \mathrm{Sinh}\left[\sqrt{2}\right]}$

Any other control than $\alpha[t] = \left(\dfrac{\mathrm{Sinh}\left[\sqrt{2}\,(-1+t)\right]}{\sqrt{2}\,\mathrm{Cosh}\left[\sqrt{2}\right] - \mathrm{Sinh}\left[\sqrt{2}\right]}\right)$

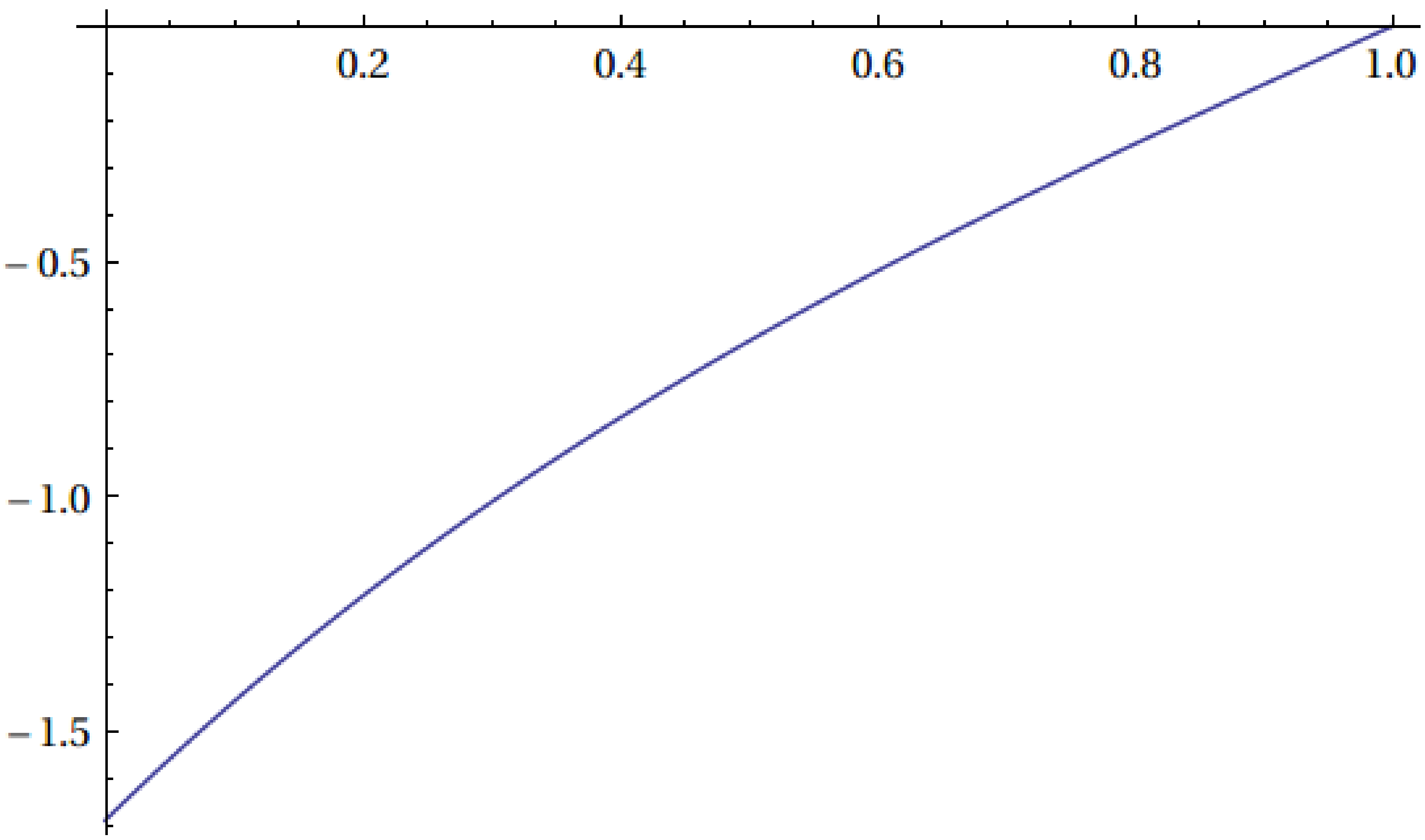
will result in

$$Q[\alpha[t]] = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) \, dt \geq 1.6895$$

Let us check that.

plot the optimal α[t] :

$$\text{Plot}\left[\frac{\text{Sinh}\left[\sqrt{2}\;(-1+t)\right]}{\sqrt{2}\;\text{Cosh}\left[\sqrt{2}\right]-\text{Sinh}\left[\sqrt{2}\right]},\ \{t,\ 0,\ 1\}\right]$$



We could try something like α[t] = a = constant (presumably negative).

Then (from x'[t] - x[t] == a) :

x[t] = A + B Exp[γ t]

```
Clear[γ]

D[ (A + B Exp[γ t]), t] - (A + B Exp[γ t]) // FullSimplify
```

$-A + B\, e^{t\,\gamma}\,(-1+\gamma)$

Therefore we want A, B, γ such that

$-A + B\, e^{t\,\gamma}\,(-1+\gamma) = a$, or $\quad -A + B\, e^{t\,\gamma}\,(-1+\gamma) - a == 0$

$\left(-A + B\, e^{t\,\gamma}\,(-1+\gamma) - a\right)$ /. $\{\gamma \to 1\}$

$-a - A$

Then $\{A = -a,\ \gamma = 1\}$

Furthermore, $x[t = 0] == 1$ and $x[t] = A + B\,\mathrm{Exp}[\gamma\, t] \Rightarrow A + B = 1$

Then $\{A = -a,\ B = 1 + a,\ \gamma = 1\}$

$\{x[t] \to A + B\,\mathrm{Exp}[\gamma\, t]\}$ /. $\{A \to -a,\ B \to 1 + a,\ \gamma \to 1\}$

$\left\{x[t] \to -a + (1+a)\, e^{t}\right\}$

So the choice $\alpha[t] = a \Rightarrow x[t] = -a + (1+a)\, e^{t}$

With this choices, the minimum of

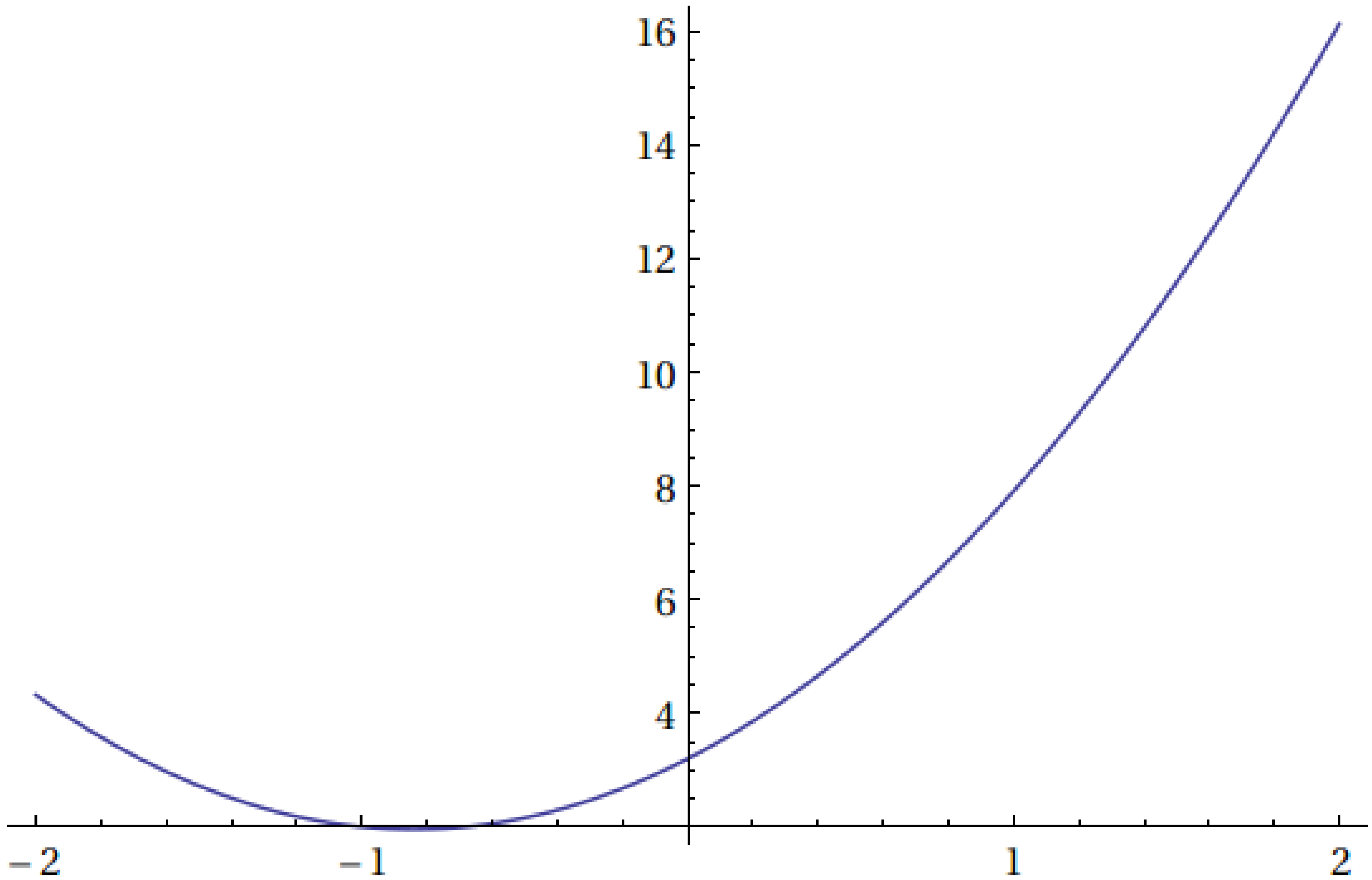
$$\int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt$$

is :

$$\int_0^1 \left(\left(-a + (1+a)\, e^{t}\right)^2 + a^2 \right) dt$$

$$\frac{1}{2}\left(-1 + 2\,a\,(-1+e)^2 + e^2 + a^2\left(7 - 4\,e + e^2\right)\right)$$

```
Plot[1/2 (-1 + 2 a (-1 + e)^2 + e^2 + a^2 (7 - 4 e + e^2)), {a, -2, 2}]
```




```
D[1/2 (-1 + 2 a (-1 + e)^2 + e^2 + a^2 (7 - 4 e + e^2)), a] // Simplify
```

$(-1 + e)^2 + a\,(7 - 4\,e + e^2)$

```
Solve[% == 0, a] // Flatten
```

$\left\{a \to -\frac{(-1 + e)^2}{7 - 4\,e + e^2}\right\}$

```
% // N
```

{a → -0.839748}

```
1/2 (-1 + 2 a (-1 + e)^2 + e^2 + a^2 (7 - 4 e + e^2)) /. %
```

1.95485

That's the minimum if we take $\alpha[t]$ = constant = -0.839748

$Q\ [\alpha[t]$ = constant = -0.839748$]$ = 1.95485

This is above the absolute real minimum obtained via optimization theory

which is (see above) 1.6895 :

$$Q[\alpha[t]] = \int_0^T \left(x[t]^2 + \alpha[t]^2\right) dt \ \geq \ Q\left[\alpha^*[t]\right] = 1.6895$$

For the general case of evolution :

$\dot{X}[t] = f[X[t], \alpha[t]] \equiv F$

let us recall our previous comments, and extend them a bit :

1. by construction, the first hamilton equation $\dot{X} = \frac{\partial H}{\partial P}$

is just a reformulation of $\dot{X}[t] = f[X[t], \alpha[t]] \equiv F$

2. The second hamilton equation $\dot{P} = -\frac{\partial H}{\partial X}$ enables us to

follow the evolution of the new auxiliary function $P[t]$

which then in turn provides $\alpha^*[t]$ (***).

3. The above program is theoretically feasible, if we knew the initial values of P[t]. Note that $P[t] = \{p_1[t], \ldots, p_n[t]\}$ contains as many functions (or Lagrange multipliers) as conditions enter in the evolution equation $\dot{X}[t] = f[X[t], \alpha[t]] \equiv F$.

4. We can denote $\{\lambda_1, \ldots, \lambda_n\} \equiv \{p_1[0], \ldots, p_n[0]\}$ and think that the optimal control $\alpha^*[t]$ can be characterized by the λ parameters. Control theory tells us that there exists a particular set of $\{\lambda_1, \ldots, \lambda_n\}$ which provides the absolute maximum of the payoff $P\left[\alpha^*[.]\right]$ (*)

5. The search for the optimum $\{\lambda_1, \ldots, \lambda_n\}^*$ values can be performed by means of a Monte Carlo optimization. We will need to numerically conduct the evolution of the set of equations

$$\dot{X} = \frac{\partial H}{\partial P}$$

$$\dot{P} = -\frac{\partial H}{\partial X}$$

where $\alpha^*[t]$, the optimal control, takes the value a which maximizes $H = H[X, P, a] = P . f + r \equiv$ control theory Hamiltonian

6. As a result, we are in the position to numerically compute the payoff $P[\alpha[.]] = \int_0^T r[x[t], \alpha[t]]\, dt + g[x[T]]$ (*)

6. As a result, we are in the position to numerically compute the

payoff $P[\alpha[.]] = \int_0^T r[x[t], \alpha[t]]\, dt + g[x[T]]$ (*)

as a function of $\{\lambda_1, ..., \lambda_n\}$: $P = P[\lambda]$ and perform a Monte Carlo search of the optimum $\{\lambda_1, ..., \lambda_n\}^*$ values by defining a probability distribution for the λ values

$$p[\lambda] = Exp[P[\lambda] / T]$$

7. Should we be interested in the minimization of a Cost function, we can define the payoff $P[\alpha[.]] = -Cost$ and then find the minimum cost performing a Monte Carlo exploration by defining a probability distribution for the λ values

$$p[\lambda] = Exp[-Cost[\lambda] / T]$$