Direct computation of optimal controls
The gradient method

Suppose that we have to optimise

$$J = \Phi(\underline{x}(T), T) + \int_0^T F(t, \underline{x}, \underline{u})\, dt$$

subject to a dynamic constraint

$$\dot{\underline{x}} = \underline{f}(t, \underline{x}, \underline{u}) \quad \underline{x}(0) = \underline{a}$$

If we introduce a new state variable x_{n+1} which satisfies

$$\dot{x}_{n+1} = F(t, \underline{x}, \underline{u}) \quad x_{n+1}(0) = 0$$

then

$$\int_0^T F(t, \underline{x}, \underline{u})\, dt = x_{n+1}(T)$$

and we can write the cost function as

$$J = \Phi(\underline{x}(T), T) + x_{n+1}(T)$$

Therefore, without loss of generality we can consider a cost function which is a function of the terminal values only.

Suppose that we have a first approximation $\underline{u}_0(t)$ for the optimal control and that we have calculated the corresponding solution $\underline{x}_0(t)$.

If we perturb the control so that $\underline{u} = \underline{u}_0 + \underline{\delta u}$, we induce a perturbation in the solution $\underline{x} = \underline{x}_0 + \underline{\delta x}$, where we have

$$\begin{aligned}
\frac{d}{dt}(\underline{x}_0 + \underline{\delta x}) &= \underline{f}(t, \underline{x}_0 + \underline{\delta x}, \underline{u}_0 + \underline{\delta u}) \\
&\doteq f(t, \underline{x}_0, \underline{u}_0) \\
&\quad + \frac{\partial \underline{f}}{\partial \underline{x}}(t, \underline{x}_0, \underline{u}_0)\, \underline{\delta x} \\
&\quad + \frac{\partial \underline{f}}{\partial \underline{u}}(t, \underline{x}_0, \underline{u}_0)\, \underline{\delta u} \\
\dot{\underline{\delta x}} &\doteq A \underline{\delta x} + B \underline{\delta u}
\end{aligned}$$

where A is the Jacobian matrix $(\partial f_i / \partial x_j)$ and B is the Jacobian matrix $(\partial f_i / \partial u_j)$; both matrices being evaluated along the first approximate solution.

The solution of this approximate linear system can be written in the form

$$\begin{aligned}
\underline{\delta x} &= \mathcal{X}(t)\mathcal{X}^{-1}(0)\underline{\delta x}(0) \\
&\quad + \int_0^t \mathcal{X}(t)\mathcal{X}^{-1}(s)B(s)\underline{\delta u}(s)\, ds \\
&= \int_0^t \mathcal{X}(t)\mathcal{X}^{-1}(s)B(s)\underline{\delta u}(s)\, ds \quad \text{since } \underline{\delta x}(0) = \underline{0} \\
\underline{\delta x}(T) &= \int_0^T \mathcal{X}(T)\mathcal{X}^{-1}(s)B(s)\underline{\delta u}(s)\, ds
\end{aligned}$$

The perturbation also alters the value of the cost function.

$$\Phi(\underline{x}(T),T) \doteq \Phi(\underline{x}_0(T),T) + \nabla\Phi(\underline{x}_0(T),T).\underline{\delta x}(T)$$

If $\underline{x}$ is in fact optimal then $\nabla\Phi(\underline{x}_0(T),T) = \underline{0}'$.
Otherwise we can improve the value of Φ by a suitable choice of $\underline{\delta x}(T)$.
For maximum effect choose $\underline{\delta x}(T)$ parallel to $\underline{g} = [\nabla\Phi(\underline{x}_0(T),T)]'$.
We need to choose $\underline{\delta u}$ so that

$$\int_0^T \mathcal{X}(T)\mathcal{X}^{-1}(s)B(s)\underline{\delta u}(s)\,ds = \epsilon \underline{g}$$

where ϵ is a scaling factor chosen to keep $\underline{\delta u}$ *small*.

ϵ is positive if we wish to maximise J and negative if we wish to minimise J.

If we set

$$\underline{\delta u}(s) = B'(s)(\mathcal{X}^{-1}(s))'\mathcal{X}(T)'\underline{c}$$

then

$$\left[\int_0^T \mathcal{X}(T)\mathcal{X}^{-1}(s)B(s)B'(s)\mathcal{X}^{-1}(s)'\mathcal{X}(t)\,ds\right]\underline{c} = \epsilon \underline{g}$$

$$\mathcal{U}(T,0)\underline{c} = \epsilon \underline{g}$$

$$\underline{c} = \epsilon \mathcal{U}^{-1}(T,0)\underline{g}$$

The matrix $[\mathcal{X}(T)\mathcal{X}^{-1}(s)]'$ is the solution of the matrix differential equation

$$\frac{dX}{ds} = -A'(s)X(s)\ ;\ X(T) = I$$

so that $\mathcal{U}$ is calculated numerically from $s = T$ to $s = 0$.

Once $\underline{\delta u}$ is calculated, we upgrade our approximate control from $\underline{u}_0$ to $\underline{u}_1 = \underline{u}_0 + \underline{\delta u}$ and repeat the process from the top.

The gradient method for discrete systems.

Consider

$$x(n+1) = f(x(n), u(n), n)$$
$$x(0) = x_0$$

For some fixed N, we want to find

$$u(0) \dots u(N-1)$$

such that $P(x(N))$ is maximised.

Suppose that we know an approximate control $U(n)$ with corresponding solution $X(n)$.

If we use instead the control

$$u(n) = U(n) + du(n)$$

we obtain the solution

$$x(n) = X(n) + dx(n)$$

For du and dx sufficiently small,

$$\begin{aligned}
x(n+1) &= X(n+1) + dx(n+1) \\
&= f(X(n) + dx(n), U(n) + du(n), n) \\
&= f(X(n), U(n), n) \\
&\quad + \frac{\partial f}{\partial x}(X(n), U(n), n)\, dx(n) \\
&\quad + \frac{\partial f}{\partial u}(X(n), U(n), n)\, du(n) \\
X(n+1) &= f(X(n), U(n), n) \\
dx(n+1) &= A(n)dx(n) + B(n)du(n)
\end{aligned}$$

where $A(n)$ and $B(n)$ are the Jacobian matrices evaluated along the approximate solution.

Solving the linear recurrence relation we have

$$\begin{aligned}
dx(1) &= A(0)dx(0) + B(0)du(0) \\
&= B(0)du(0) \qquad \text{since} \quad dx(0) = 0 \\
dx(2) &= A(1)dx(1) + B(1)du(1) \\
&= A(1)B(0)du(0) + B(1)du(1) \\
dx(N) &= A(N-1)\dots A(1)B(0)du(0) \\
&\quad + A(N-1)\dots A(2)B(1)du(1) \\
&\quad + \dots \\
&\quad + A(N-1)B(N-2)du(N-2) \\
&\quad + B(N-1)du(N-1)
\end{aligned}$$

Our objective function

$$\begin{aligned} P(x(N)) &= P(X(N) + dx(N)) \\ &= P(X(N) + \nabla P(X(N)).dx(N) \\ dP &= \nabla P(X(N)).dx(N) \end{aligned}$$

We want to choose $dx(N)$ so that dP is maximised. Therefore we choose $dx(N)$ as a multiple of $(\nabla P)^T (= g(N))$.

This gives

$$\begin{aligned} dP &= cg^T.dx(N) \\ &= c[g^T A(N-1)\dots A(1)B(0)du(0) \\ &\quad + \dots \\ &\quad + g^T B(N-1)du(N-1)] \end{aligned}$$

and we ensure that all the contributions to this are positive by choosing

$$\begin{aligned} du(0) &= B^T(0)A^T(1)\dots A^T(N-1)g(N) \\ du(1) &= B^T(1)A^T(2)\dots A^T(N-1)g(N) \\ &\dots \\ du(N-1) &= B^T(N-1)g(N) \end{aligned}$$