which must hold for all $x \in \mathbf{R}^n$. This means that the matrices P_n must satisfy the discrete-time Riccati equation

$$\begin{aligned} P_N &= Q_0 \\ P_n &= Q + A^T(P_{n+1} - P_{n+1}B(R + B^T P_{n+1}B)^{-1}B^T P_{n+1})A \end{aligned} \tag{2.5}$$

for $n = N-1, N-2, \ldots, 0$. Note that $P_n \geq 0$ for all n, which implies that the inverse in (2.5) is well defined. To see that $P_n \geq 0$ we notice that $P_N \geq 0$ because $Q_0 \geq 0$. Furthermore, the minimum in (2.4) must be positive for $n = N-1$ because P_N, Q and R are all positive semidefinite. Induction proves the result.

To summarize, we have that the optimal cost-to-go and the optimal feedback control law are

$$\begin{aligned} J(n,x) &= x^T P_n x \\ u_n^* &= \mu(n,x) = -(R + B^T P_{n+1} B)^{-1} B^T P_{n+1} A x \end{aligned}$$

where P_n is the solution to the Riccati equation in (2.5).

2.2 A Discrete Version of PMP

Consider the following discrete time optimal control problem

$$\min \phi(x_N) + \sum_{k=0}^{N-1} f_0(k, x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = f(k, x_k, u_k), \\ x_0 \text{ is given, } G(x_N) = 0 \end{cases} \tag{2.6}$$

where $G(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_p(x) \end{bmatrix}$ satisfies the usual regularity assumption, i.e. the gradients $\nabla g_k(x)$ are linearly independent. This is a special case of (2.2) in which $X = \mathbf{R}^n$ and $U = \mathbf{R}^m$. The dynamical programming approach to solving such problems is characterized by the following properties

- Feedback solutions are obtained. This means that we know the optimal control value for every position of the state vector x. This gives robustness to the closed loop system in the following sense: If the solution is perturbed by a disturbance then the controller still knows the optimal action.

- The solution is obtained using backwards recursion, which can be computationally demanding. One way to understand this is that we compute the optimal control value for every possible system state. What we win in robustness we loose in computational complexity.

- Sufficient condition.

We next use the Lagrange multiplier rule (also known as the Karush-Kuhn-Tucker conditions (KKT), or the first order optimality conditions) to obtain necessary conditions for optimality. The resulting conditions are the discrete version of the so-called Pontryagin minimization principle that we will study later in the course.

We recall from the optimization courses (the KKT conditions)

First order necessary condition (KKT): Suppose x^* is a (locally) optimal solution of

$$\min f(x) \quad \text{subject to} \quad G(x) = 0$$

where $f : \mathbf{R}^n \to \mathbf{R}$ and $G : \mathbf{R}^n \to \mathbf{R}^p$ are continuously differentiable and the constraint set is regular, i.e., the gradients $\nabla g_k(x)$ are linearly independent. Then there exists a vector of *Lagrange multipliers* $\lambda \in \mathbf{R}^p$ such that

(*i*) $G(x^*) = 0$

(*ii*) $\nabla_x l(x^*, \lambda) = 0$, where $l(x, \lambda) = f(x) + \lambda^T G(x)$ is the *Lagrangian*.

We can use it to derive the following result.

Proposition 1. *Let* $\{u_k^*\}_{k=1}^{N-1}$ *be an optimal control for (2.6) and let* $\{x_k^*\}_{k=0}^{N}$ *be the corresponding trajectory. Then there exists an adjoint variable (Lagrange multiplier)* $\{\lambda_k\}_{k=1}^{N}$ *such that*

(*i*) *(adjoint equation)*

$$\lambda_k = \frac{\partial H}{\partial x}(k, x_k^*, u_k^*, \lambda_{k+1}), \quad k = 1, \ldots, N-1$$

(*ii*) *("pointwise minimization")*

$$\frac{\partial H}{\partial u}(k, x_k^*, u_k^*, \lambda_{k+1}) = 0, \quad k = 0, 1, \ldots, N-1$$

(*iii*) *(Boundary condition)*

$$\lambda_N = \frac{\partial \phi}{\partial x}(x_N^*) + G_x(x_N^*)^T \nu$$

for some $\nu \in \boldsymbol{R}^p$.

where the Hamiltonian is

$$H(k, x, u, \lambda) = f_0(k, x, u) + \lambda^T f(k, x, u)$$

Proof. Let

$$
\begin{aligned}
z &= \begin{bmatrix} x_1^T & \dots & x_N^T & u_0^T & \dots & u_{N-1}^T \end{bmatrix}^T \\
\mathcal{F}(z) &= \phi(x_N) + \sum_{k=0}^{N-1} f_0(k, x_k, u_k) \\
\mathcal{G}(z) &= \begin{bmatrix} f(0, x_0, u_0) - x_1 \\ \vdots \\ f(N-1, x_{N-1}, u_{N-1}) - x_N \\ G(x_N) \end{bmatrix}
\end{aligned}
$$

The Lagrange multiplier rule says that a necessary condition for optimality of

$$\min \mathcal{F}(z) \quad \text{subject to} \quad \mathcal{G}(z) = 0$$

is that there exists a Lagrange multiplier $\hat{\lambda}$ such that

$$\frac{\partial l}{\partial z}(z^*, \hat{\lambda}) = 0 \quad \text{where} \quad l(z, \hat{\lambda}) = \mathcal{F}(z) + \hat{\lambda}^T \mathcal{G}(z)$$

In our problem the Lagrange multiplier vector is $\hat{\lambda} = \begin{bmatrix} \lambda^T & \nu^T \end{bmatrix}^T$. We get

$$
\begin{aligned}
\frac{\partial l}{\partial x_k}(z^*) &= \frac{\partial f_0}{\partial x}(k, x_k^*, u_k^*) + \lambda_{k+1}^T \frac{\partial f}{\partial x}(k, x_k^*, u_k^*) - \lambda_k, \quad k = 1, \dots, N-1 \\
\frac{\partial l}{\partial x_N}(z^*) &= \frac{\partial \phi}{\partial x}(x_N^*) - \lambda_N + G_x(x_N^*)^T \nu \\
\frac{\partial l}{\partial u}(z^*) &= \frac{\partial f_0}{\partial u}(k, x_k^*, u_k^*) + \lambda_{k+1}^T \frac{\partial f}{\partial u}(k, x_k^*, u_k^*), \quad k = 1, \dots, N-1
\end{aligned}
$$

Hence, the condition $\frac{\partial l}{\partial z}(z^*, \hat{\lambda}) = 0$ together with the definition of the Hamiltonian function $H(k, x_k, u_k, \lambda_{k+1})$ proves the proposition. □

The propostion is often used in the following way

1. Define the Hamiltonian: $H(k, x, u, \lambda) = f_0(k, x, u) + \lambda^T f(k, x, u)$
2. Perform pointwise minimization, i.e. find a function $\mu(k, x, \lambda)$ such that $\frac{\partial H}{\partial u}(k, x, u, \lambda) = 0$. Hence the candidate optimal control is $u_k^* = \mu(k, x_k^*, \lambda_k)$.
3. Solve the two point boundary value problem (TPBVP)

$$
\begin{aligned}
x_{k+1} &= \frac{\partial H}{\partial \lambda}(k, x_k, \mu(k, x_k, \lambda_{k+1}), \lambda_{k+1}) = f(k, x_k, \mu(k, x_k, \lambda_{k+1})), \quad G(x_N) = 0 \\
\lambda_k &= \frac{\partial H}{\partial x}(k, x_k, \mu(k, x_k, \lambda_{k+1}), \lambda_{k+1}), \quad \lambda_N = \frac{\partial \phi}{\partial x}(x_N) + G_x(x_N)^T \nu
\end{aligned}
$$

We call this a two point boundary value problem because the only unknown to determine are λ_0 and x_N. Once they are known then all other state and adjoint variables can be computed from the recursive equations. It is interesting to note that the nonlinear program in (2.6) has a lot of structure that can be exploited.

The PMP approach is characterized by the following properties.

- It results in an open loop control program. This mean that the optimal solution is only known for a particular initial condition x_0. If the solution is perturbed from the optimal by a disturbance then the optimal control may no longer be effective. The resulting system is therefore more sensitive to disturbances.
- It is generally easier to compute.
- It gives only a necessary condition for optimality.

2.3 Infinite Time Horizon Optimization

Let us consider multistage decision problems over an infinite time horizon. We consider the following general form of such problems

$$\min \Sigma_{k=0}^{\infty} f_0(x_k, u_k) \quad \text{subj. to} \quad \begin{cases} x_{k+1} = f(x_k, u_k) \\ x_0 \text{ given} \\ u_k \in U(x_k) \end{cases} \tag{2.7}$$

In order for the cost to be finite we need that $f_0(x_k, u_k) \to 0$ as $k \to \infty$. An interpretation is that (2.7) models problems where convergence to some particular set of values is desired. In our discussion we will assume that the state vector converges to zero.

The following assumptions are made

Assumption 1. We assume (w.l.o.g) that $0 \in X$, $U(0) = \{0\}$, $f(0,0) = 0$ and $f_0(0,0)$.

The assumption implies that zero is an equilibrium point of the discrete dynamics

$$x_{k+1} = f(x_k, u_k)$$

This means that if $(x_0, u_0) = (0,0)$ then the zero control $u_k = 0$, $\forall k$ implies that $x_k = 0$, $\forall k$. In order to obtain the simplest possible result we will assume that the cost function and therefore the *value function* (the optimal cost of (2.7)) are positive definite.

Definition 1. A function $V : \mathbf{R}^n \to \mathbf{R}$ is called *strictly positive definite*[2] if $V(0) = 0$ and there exists $\epsilon > 0$ such that $V(x) \geq \epsilon \|x\|^2$ for all $x \in \mathbf{R}^n$.

Example 8. A quadratic form $V(x) = x^T P x$, where $P = P^T$, is strictly positive definite if $P > 0$, i.e., if all eigenvalues of P are positive.

Assumption 2. We assume that f_0 is strictly positive definite, i.e. there exists $\epsilon > 0$ such that $f_0(x, u) \geq \epsilon(|x_k|^2 + |u_k|^2)$.

Let us now define the optimal function (value function) corresponding to (2.7)

$$J^*(x_0) = \min_{u_k \in U} \sum_{k=0}^{\infty} f_0(x_k, u_k)$$

The value function is independent of time since the dynamics and cost function of (2.7) both are independent of time (the stage index k).

Theorem 2. *Suppose Assumption 1 and Assumption 2 hold. If there exists a strictly positive definite, function $V : X \to \boldsymbol{R}^+$ that satisfies the Bellman equation*

$$V(x) = \min_{u \in U(x)} \{f_0(x, u) + V(f(x, u))\} \tag{2.8}$$

then

(*a*) $V(x) = J^*(x)$

(*b*) $u^* = \mu(x) = \text{argmin}_{u \in U(x)} \{f_0(x, u) + V(f(x, u))\}$ *is an optimal feedback control that results in a globally convergent closed loop system, i.e. for any $x_0 \in X$ the optimal solution $(x_k, \mu(x_k)) \to 0$.*

Remark 2. The assumptions are stronger than necessary but it simplifies the proof.

Proof. Sketch: We first prove that $u_k = \mu(x_k)$ gives a globally convergent system. From the Bellman equation we get

$$\begin{aligned} \sum_{k=0}^{N} f_0(x_k, \mu(x_k)) &= \sum_{k=0}^{N} V(x_k) - V(x_{k+1}) \\ &= V(x_0) - V(x_{N+1}) \end{aligned}$$

where we use $x_{k+1} = f(x_k, \mu(x_k))$. Since $V(x) \geq 0$ and N was arbitrary, we have

$$\lim_{N \to \infty} \sum_{k=0}^{N} f_0(x_k, u_k) \leq V(x_0) \tag{2.9}$$

[2]This definition of strictly positive definite is much stronger than normal. It is used to simplify the understanding of this section.

Hence, since $f_0(x_k, \mu(x_k))$ is strictly positive definite it follows that $f_0(x_k, \mu(x_k)) \to 0$ as $k \to \infty$. Otherwise the sum would be unbounded, which violates (2.9). This also implies that $(x_k, \mu(x_k)) \to (0, 0)$.

We have now proved that $u = \mu(x)$ is "stabilizing" in the sense that the closed loop state vector converges to zero. We will next see that it also gives the minimal cost. Consider an arbitrary control sequence $\{u_k\}_{k=0}^{\infty}$, which results in a convergent solution. We have

$$\begin{aligned} \lim_{N\to\infty} \sum_{k=0}^{N} f_0(x_k, u_k) &\geq \lim_{N\to\infty} \sum_{k=0}^{N} V(x_k) - V(x_{k+1}) \\ &= V(x_0) - \lim_{N\to\infty} V(x_N) = V(x_0) \end{aligned}$$

where we used that $x_N \to 0 \Rightarrow V(x_N) \to 0$. Since the first inequality becomes an equality when $u_k = \mu(x_k)$, we get

$$V(x_0) = \sum_{k=0}^{\infty} f_0(x_k, \mu(x_k)) \leq \sum_{k=0}^{\infty} f_0(x_k, u_k)$$

This proves the optimality. □