Comenius University in Bratislava

Faculty of Mathematics, Physics and Informatics

Maximum Principle for Infinite Horizon Discrete Time Optimal Control Problems

Dissertation Thesis

Study program: 9.1.9. Applied Mathematics

Supervising institution: Department of Applied Mathematics and Statistics

Supervisor: prof. RNDr. Pavel Brunovský, DrSc.

Bratislava 2016

Mgr. Mária Holecyová

Univerzita Komenského v Bratislave

Fakulta matematiky, fyziky a informatiky

Princíp maxima pre diskrétne úlohy optimálneho riadenia na nekonečnom horizonte

Dizertačná práca

Študijný program: 9.1.9. Aplikovaná matematika

Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky

Školiteľ: prof. RNDr. Pavel Brunovský, DrSc.

Bratislava 2016

Mgr. Mária Holecyová

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor prof. RNDr. Pavel Brunovský, DrSc. for the continuous support of my PhD study, for his patience, motivation and immense knowledge. Undoubtedly, the thesis would not have been completed without his guidance and unwavering belief that 'it certainly has to work this way we just have to prove it'.

And I would like to thank to my family and friends whose encouragmenet and understading helped me accomplish this study as well.

Thank you all.

Abstract

Mgr. Mária Holecyová: *Maximum Principle for Infinite Horizon Discrete Time Optimal Control Problems* [Dissertation Thesis] Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, supervisor: prof. RNDr. Pavel Brunovský, DrSc., Bratislava, 2016, 76p.

The aim of this thesis is a method of deriving necessary conditions of the Potryagin maximum principle type for infinite-horizon discrete-time optimal control problems with discount. Due to the discounted objective function, control and state variables are considered to be bounded sequences. We employ the tools of functional analysis and properties of linear difference systems.

Firstly, we prove Fréchet differentiability of the objective function which allows us to carry out a standard method of obtaining necessary conditions of optimality of variational type. Then we apply the closed range theorem and formulate maximum principle in functional form with adjoint variable from the space $(\ell_\infty)^* = \ell_1 \oplus \ell_s$. Then we show that it can be rewritten to the standard form of Potryagin maximum principle for adjoint variable belonging to ℓ_1 .

The most significant results are conditions under which the assumptions of the closed range theorem are satisfied. For a problem with linear dynamics we require that the matrix A has no eigenvalues on the unit circle and in case of general dynamics we formulate exponential dichotomy as an assumption. We present special cases in which exponential dichotomy can be effectively verified. In addition, on a simple example we show that without exponential dichotomy the assumption of closed range probably may not hold.

Keywords: optimal control, discrete time, infinite horizon, Pontryagin maximum principle, closed range theorem, ℓ_∞

Abstrakt

Mgr. Mária Holecyová: *Princíp maxima pre diskrétne úlohy optimálneho riadenia na nekonečnom horizonte* [Dizertačná práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, školiteľ: prof. RNDr. Pavel Brunovský, DrSc., Bratislava, 2016, 76s.

Cieľom tejto práce je metóda pre odvodenie nutných podmienok optimality vo forme Potrjaginovho princípu maxima pre diskrétne úlohy optimálneho programovania na nekonečnom horizonte s diskontom. Predpokladáme, že stavová aj riadiaca premenná sú ohraničené postupnosti, využívame prostriedky funkcionálnej analýzy a vlastnosti lineárnych diferenčných rovníc.

Dokazujeme, že účelová funkcia je Fréchetovsky diferencovateľná, čo nám umožňuje uplatniť štandardnú metódu pre formulovanie nutných podmienok optimality variačného typu. Ďalej využívame vetu o uzavretom obraze a formulujeme princíp maxima vo funkcionálnom tvare s adjungovanou premennou z priestoru $(\ell_\infty)^* = \ell_1 \oplus \ell_s$. Následne ukazujeme, že ho možno prepísať do štandardnej formy Pontrjaginovho princípu maxima pre adjungovanú premennú z priestoru ℓ_1.

Najvýznamnejším výsledkom nášho výskumu je formulácia podmienok, ktoré zaručujú splnenie predpokladov vety o uzavretom obraze. V prípade lineárnej autonómnej dynamiky stačí, aby matica A nemala vlastné čísla ležiace na jednotkovej kružnici, v prípade všeobecnej dynamiky formulujeme ako predpoklad exponenciálnu dichotómiu.

Na záver sa venujeme špeciálnym prípadom, v ktorých je možné overiť existenciu exponenciálnej dichotómie a na jednoduchom prípade ukazujeme, že predpoklad uzavretého obrazu nemusí bez exponenciálnej dichotómie platiť.

Kľúčové slová: optimálne riadenie, diskrétny čas, nekonečný horizont, Pontrjaginov princíp maxima, veta o uzavretom obraze, ℓ_∞

Preface

"If every instrument could accomplish its own work, obeying or anticipating the will of others... the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not need servants, nor masters slaves." *Aristotle, Politcs*

The motivation for solving optimal control problems goes from the simplest mechanisms we manipulate in everyday life to the most sophisticated ones in various fields of science. Although the discipline is quite young, it is considered to be invented only in 1950's by Pontryagin, we already have effective tools to solve a variety of optimal control problems and with advancing progress more scientific applications from different fields are appearing.

Currently, we distinguish between two main tools - Bellman's dynamical programming originally developed for discrete time models and Potryagin maximum principle developed for continuous time models. Our aim is to combine the Potryagin variational approach with discrete-time infinite-horizon models by applying the closed range theorem. In the thesis we refer to two articles closely associated with this topic, but it is the diploma thesis that stands at the beginning of our research. This is where my supervisor prof. Pavel Brunovský first came with the idea of employing the closed range theorem.

Only thanks to his deep knowledge in various mathematical fields, his creativity and unceasing enthusiasm we managed to combine optimal control theory with functional analysis and properties of linear difference equations which resulted in this thesis.

Contents

Introduction

In economic optimal control models [2] as well as in physics [3], engineering [4] and many other fields it is often impossible to predict the length of the time horizon. Therefore an objective function is formulated on infinite horizon and especially in economic models it is discounted. The discount ensures that the effect of the solutions to the objective function decreases with passing time which solves the dilemma of setting the length of the horizon as well as the final state.

We consider discrete-time problems on infinite horizon with discounted objective function and we focus on establishing neccesary conditions of optimality in the spirit of Pontryagin maximum principle which was originally devoloped for continuous-time models.

While for the continuous-time setting Pontryagin maximum principle can be easily adapted for a wide class of problems, this is not the case of the discrete-time problems unless extra convexity conditions are imposed. So instead of the maximum condition we strive for a necessary condition of this maximum with less restrictive assumptions.

The current research on this topic is not rich. We found only two articles closely related to our problem. In the first Blot and Chebbi [5] solved it as limit case of finite horizon problem. In the continous framework, the extension from finite to infinite horizon can be obtained without any restrictions due to its invertible dynamics. However, in discrete time invertibility is not ensured and therefore it has to be formulated as an additional condition. Blot and Hayek [6] managed to formulate less strict conditions directly via functional analysis. We adapt their approach, but while their results are based on Ioffe-Tihomirov theorem, we employ the closed range theorem. This idea first appeared in diploma thesis by Beran [7] but it was developed for non-discounted objective functions and the research on conditions under which the theorem can be employed was incomplete.

The thesis is organized as follows. In the first chapter, we introduce the optimal control problem which is considered thoughout the whole thesis as well as the closed range theorem. The second chapter is devoted to an overview of previous results regarding Potryagin principle from discrete-time infinite-horizon view. It also describes a dual space of the space of all bounded sequences ℓ_∞ and refers to the literature dealing with its singular component. In the next chapters we prove that the considered objective function is Fréchet differentiable and then we propose a method of obtaining the maximum principle for problems with both linear autonomous and general dynamics. We apply the closed range theorem and assume that the respective operator has closed range. In the final chapter, we formulate condition under which this is the case by introducing exponential dichotomy for difference equations.

Chapter 1

Problem formulation

Our motivation comes mainly from macroeconomic optimal growth models [2], [8], [9], [10] which often cannot predict the length of the time horizon, but assume that it is large. Therefore a discount is added to the objective funtion and it is maximized on infinite horizon. In this case, we naturally expect that response $\{x_t\}_{t=0}^{\infty}$ and control $\{u_t\}_{t=0}^{\infty}$ are bounded sequences, but do not vanish in infinity necessarily. Hence, the problem we consider has the following form:

$$J(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{T} \delta^t f_t^0(x_t, u_t) \to \max \tag{1.1}$$

$$x_{t+1} = F_t(x_t, u_t) \quad \text{for all } t \in \mathbb{N}_0 \tag{1.2}$$

$$x_0 = \bar{x}, \tag{1.3}$$

where $\bar{x}$ and the discount $\delta \in (0, 1)$ are given, $x_t \in X \subset \mathbb{R}^n$, $u_t \in U \subset \mathbb{R}^m$, U open. We denote $\mathbf{x} = \{x_t\}_{t=0}^{\infty}$, $\mathbf{u} = \{u_t\}_{t=0}^{\infty}$ and assume $f_t^0 \in C^1(X \times U, \mathbb{R})$ for all $t \in \mathbb{N}_0$ $F_t \in C^1(X \times U, X)$ for all $t \in \mathbb{N}_0$. We call J objective function, F_t dynamics. x_t state variable and u_t control variable. We assume $(\mathbf{x}, \mathbf{u}) \in \ell_\infty^n \times \ell_\infty^m$. And we put $T = \infty$.

In case of finite horizon ($T < \infty$) this is a standard problem of nonlinear programming and it yields the following necessary condition of optimality [11], where a regularity condition has to be fulfilled.

Definition 1.1. (**Regularity condition**) Denote $I_t(u_t)$ as the set of all indices $k \in \{1, ..., m_t\}$ for which $p_t^k = 0$, i.e. the set of indices for which the k^{th} component of the constraint p_t is active. The optimal control problem (1.1) - (1.3) with $T \in \mathbb{N}$ *fulfills the regularity condition* in $\hat{\mathbf{u}}$, if for all $t \in \mathbb{N}_0$ the vectors

$$\frac{\partial p_t^k}{\partial u_t}(\hat{u}_t) \quad \text{for all } k \in I_t(\hat{u}_t)$$

are lineary independent.

Theorem 1.1. *(**Necessary conditions of optimality, pseudo-Pontryagin maximum principle**) Let $\hat{\mathbf{u}}$ be an optimal control for the problem (1.1) - (1.3) with $T < \infty$ and let $\hat{\mathbf{x}}$ be its response. Let the regularity condition be fulfilled in $\hat{\mathbf{u}}$ (Definition 1.1). Then there exists $\psi_0 \geq 0$, an adjoint variable $\boldsymbol{\psi} = \{\psi_t\}_{t+1}^{\infty}$, a vector $\chi \in \mathbb{R}^l$ and vectors $\lambda_i \in \mathbb{R}^{m_t}$ for all $t \in \{0, ..., T-1\}$ such that $(\psi_0, \chi) \neq 0$ and for all $t \in \{0, ..., T-1\}$ the variation condition*

$$\psi^0 \frac{\partial f_t^0}{\partial u_t}(\hat{x}_t, \hat{u}_t) + \psi_{t+1}^T \frac{\partial f_t}{\partial u_t}(\hat{x}_t, \hat{u}_t) + \lambda_t^T \frac{\partial p_t}{\partial u_t}(\hat{u}_t) = 0 \tag{1.4}$$

holds and the complementarity condition

$$\lambda_t^T p_t(\hat{u}_t) = 0, for\ \lambda_t \leq 0, \tag{1.5}$$

holds, as well. The adjoint variables and χ solve the adjoint equation

$$\psi_t = \left(\frac{\partial f_t^0}{\partial x_t}(\hat{x}_t, \hat{u}_t)\right)^T + \left(\frac{\partial f_t}{\partial x_t}(\hat{x}_t, \hat{u}_t)\right)^T \psi_{t+1} \quad for\ all\ t, \tag{1.6}$$

together with the transversality condition

$$\psi_k = \left(\frac{\partial g}{\partial x_t}(\hat{x}_k)\right)^T \chi. \tag{1.7}$$

In [11] this is called pseudo-Pontryagin maximum principle. Notice that variation condition (1.4) and complementarity condition (1.5) are in fact a necessary condition of maximum of $\psi^0 f_t^0(\hat{x}_t, u_t) + \psi_{t+1} F_t(\hat{x}_t, u_t)$, i.e. they can be replaced

$$\psi^0 f_t^0(\hat{x}_t, \hat{u}_t) + \psi_{t+1} F_t(\hat{x}_t, \hat{u}_t) = \max_{u_t \in U_t} (\psi^0 f_t^0(\hat{x}_t, u_t) + \psi_{t+1} F_t(\hat{x}_t, u_t)) \quad \text{for all } t. \tag{1.8}$$

However, in order that the equation holds futher assumptions are needed.

Theorem 1.2. ***(Pontryagin maximum principle for discrete time problems)*** *Let the conditions of Theorem 1.1 be fulfilled. Furthermore, assume that for all* $t = \{0, ..., T-1\}$

(i) the functions F_t *are linear in* u_t*,*

(ii) the functions f_t^0 *are concave in* u_t*,*

(iii) the set U *is convex,*

Then there exists vector $\chi \in \mathbb{R}^l$*,* $\psi_0 \geq 0$*, a sequence of adjoint variables* $\boldsymbol{\psi} = \{\psi_t\}_{t=1}^T$ *that are a solution of the adjoint equation (1.6) and transversality condition (1.7) such that*

$$\psi^0 f_t^0(\hat{x}_t, \hat{u}_t) + \psi_{t+1} F_t(\hat{x}_t, \hat{u}_t) = \max_{u_t \in U_t} (\psi^0 f_t^0(\hat{x}_t, u_t) + \psi_{t+1} F_t(\hat{x}_t, u_t)) \quad \textit{for all } t. \tag{1.9}$$

Pontryagin maximum principle was originally developed for continuous time problems where it can be easily adapted for infinite horizon problems. However, this is not the case of discrete-time problems. Convexity condition might significantly reduce the class of solvable problems. So instead of the maximum condition we strive for a necessary condition of maximum with less restrictive assumptions, i.e. pseudo-Pontryagin maximum principle. For this purpose, we adapt the approach by Beran [7] who studied infinite-horizon discrete-time problems without discount. Whereas the natural space for the controls/responses in the case without discount is ℓ_1, for the problem with discount it is

ℓ_∞ which has complicated dual space $(\ell_\infty)^* = \ell_1 \oplus \ell_s$, because the space ℓ_s cannot be represented by sequences and thus it requires another approach.

The central role in our research plays the closed range theorem.

Theorem 1.3. ***Closed range theorem***

Let X, Y be Banach spaces and T a closed linear operator from X to Y. Then the following propositions are equivalent:

1. $\mathcal{R}(T)$ *is closed*
2. $\mathcal{R}(T^*)$ *is closed*
3. $\mathcal{R}(T) = \mathcal{N}(T^*)^\perp = \{y \in Y : \langle y^*, y\rangle = 0 \text{ for all } y^* \in \mathcal{N}(T^*)\}$
4. $\mathcal{R}(T^*) = \mathcal{N}(T)^\perp = \{x^* \in X^* : \langle x^*, x\rangle = 0 \text{ for all } x \in \mathcal{N}(T)\}$.

Proof. The proof can be found in [12]. □

We will use only its reduced form $1 \Rightarrow 4$. It allows us to establish pseudo-Potryagin principle however condition of closed range has to satisfied. Therefore we develop the theory of exponential dichotomy for linear difference equations that is studied in Coppel [14], Palmer [13] and derive conditions under which the pseudo-Potryagin principle holds.

Chapter 2

Previous results

The aim of this chapter is to provide a brief review of the literature associated with our research. It is devided into two sections. Firstly, we summarize current knowledge about discrete time optimal control problems on infinite horizon. The second part describes methods which are being employed to deal with dual space of ℓ_∞.

2.1 Optimal Control Problems

Nowadays there are basically two solution methods in the optimal control theory - dynamic programming developed by Bellman and Pontryagin maximum principle. Dynamic programming was originally developed for discrete-time optimal control problems, while Pontryagin maximum principle was derived to solve continuous-time problems. Later, the relation between the two methods has been explained.

In this work we focus on the Pontryagin maximum principle from the infinite-horizon discrete-time point of view. Compared to continuous time problems, in the discrete time case Pontryagin maximum principle as a necessary condition of optimality requires extra conditions as it shown in Chapter 1. Unless certain convexity conditions are imposed, a necessary conditi on of maximum (in the sense of Theorem 1.1) is claimed instead of a maximum condition (in the sense of Theorem 1.2). Hence, there are two different approaches, the first one focuses on establishing Pontryagin maximum principle for problems that satisfy additional convexity assumptions and the second approach derive so-called pseudo-Pontryagin maximum principle that requires less restrictive conditions, but it is weaker. In the thesis, we focus on the second approach.

A short section dealing with the infinite horizon discrete time problems is in Pontryagin et. al [1]. It is studied as a limit case of the finite horizon problem for $T \to \infty$. The effects of the variations are transferred to a fixed time independent of t. Whereas in the finite horizon the time can be chosen as the terminal one, in the infinite horizon case this is impossible. Consequently, the effects of the variations have to be transferred backwards. This is possible in continuous-time models due to its invertible dynamics. However, it can not be applied on discrete-time problems unless extra conditions are imposed. Boltyanskii in [15] initiated a systematic study of the discrete time framework, but mostly concerning finite horizon. He emphasizes the differences between the discrete-time setting and the continuous-time setting for the Pontryagin principle and gives the first steps for a rigorous treatment of the problem.

Later, discrete-time problems on infinite horizon were studied in McKenzie [16], Michel [19], Peleg-Ryder [20], all of them considering the concave case of optimal control problem. Their results are extended in the paper by Blot and Chebbi [5], where the authors avoid concavity assumptions and establish Pontryagin maximum principle. They solve the problem with general dynamics and without discount, that is in our notation

$$J(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{\infty} f_t^0(x_t, u_t) \to \max \tag{2.1}$$

$$x_{t+1} = F_t(x_t, u_t), \ \ t \in \mathbb{N}_0 \tag{2.2}$$

$$x_0 = \bar{x}, \tag{2.3}$$

where $x_t \in X \subset \mathbb{R}^n$, $u_t \in U \subset \mathbb{R}^m$, $\mathbf{x} = \{x_t\}_{t=0}^{\infty}$, $\mathbf{u} = \{u_t\}_{t=0}^{\infty}$, the initial state $\bar{x}$ is given and $f_t^0 : X \times U \to \mathbb{R}$ and $F_t : X \times U \to X$.

They also consider two more problems where (2.1) is replaced by the following conditions

(i)

$$\liminf_{T\to\infty}\left(\sum_{t=0}^{T} f_t^0(\hat{x}_t,\hat{u}_t)-\sum_{t=0}^{T} f_t^0(x_t,u_t)\right)\geq 0$$

(ii)

$$\limsup_{T\to\infty}\left(\sum_{t=0}^{T} f_t^0(\hat{x}_t,\hat{u}_t)-\sum_{t=0}^{T} f_t^0(x_t,u_t)\right)\geq 0$$

The problems are solved in three steps consisting of reduction to a finite horizon problem, solving the latter and extension to infinite horizon. Next, they formulate and prove three theorems of the infinite-horizon Pontryagin principle kind. According to the first one, following assumptions have to be fullfilled:

1. for all $t \in \mathbb{N}$ f_t^0 is Lipshitz continuous[1] near the solution $(\hat{x}_t, \hat{u}_t)$ and is Clarke-regular[2] at $(\hat{x}_t, \hat{u}_t)$,
2. for all $t \in \mathbb{N}$ F_t is differentiable at $(\hat{x}_t, \hat{u}_t)$,
3. U_t is closed and regular at $\hat{u}_t$,
4. for all $t \in \mathbb{N}$ $D_{x_t}F_t(\hat{x}_t, \hat{u}_t)$ is invertible.

In the next theorem, the first condition is replaced by the assumption of strictly differentiability of f_t^0 for all $t \in \mathbb{N}$. The last Pontryagin principle assumes that

1. for all $t \in \mathbb{N}$ f_t^0 and F_t are partially differentiable at $(\hat{x}_t, \hat{u}_t)$ with respect to the first vector variable,
2. X is an open convex subset,
3. Michel condition: for all $t \in \mathbb{N}$ co $A_t(\hat{x}_t, \hat{x}_{t+1}) \subset B_t(\hat{x}_t, \hat{x}_{t+1})$ is fulfilled, where co denotes the convex hull. When $(x_t, x_{t+1}) \in X \times X$, $A_t(x_t, x_{t+1})$ is the set of the points $(\lambda, y) \in \mathbb{R} \times \mathbb{R}^n$, for which there exists $u \in U$ such that $\lambda \leq F_t^0(x_t, u)$ and

[1] Given two metric spaces X, Y, a function $f : X \to Y$ is called *Lipschitz continuous* if there exists a real constant $K \geq 0$ such that, for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) \leq K d_X(x_1, x_2)$.

[2] For the definition see e.g. [22].

$y = F_t(x_t, u) - x_{t+1}$. $B_t(x_t, x_{t+1})$ is the set of $(\lambda, y) \in \mathbb{R} \times \mathbb{R}^n$ for which there exist $(u, v) \in U \times \mathbb{R}^n$ such that $\lambda \leq f_t^0(x_t, u)$ and $v^h y^h = F_t^h(x_t, u) - x_{t+1}^h$, for every $h = 1, ..., n$.

Hence, except of the last theorem, regularity of $D_{x_t} F_t(\hat{x}_t, \hat{u}_t)$ is required that significantly reduces the framework of solvable problems.

Blot and Hayek in [6] build on these results trying to avoid the regularity condition, they considered the space of all bounded sequences ℓ_∞. They establish Pontryagin maximum principle using analysis in Banach spaces instead of reduction to finite-horizon problems. They also manage to formulate sufficient conditions of optimality. In our notation, they solve the following problem with general dynamics

$$\begin{aligned} J(\mathbf{x}, \mathbf{u}) &= \sum_{t=0}^{\infty} \delta^t f^0(x_t, u_t) \to \max \\ x_{t+1} &= F_t(x_t, u_t), \;\; t \in \mathbb{N}_0 \\ x_0 &= \bar{x}, \end{aligned} \tag{2.4}$$

Compared to the problem (2.1) - (2.3) from [5], they add discount $\delta \in (0, 1)$ to the objective function and assume that $(\mathbf{x}, \mathbf{u}) \in \ell_\infty^{n+m}$.

They obtain neccesary optimality conditions assuming that

1. for all $\mathbf{u} \in \ell_\infty^m$, the mapping $x \to f^0(x, u)$ is of class C^1 on X and for all $t \in \mathbb{N}$, the mapping $x \to F_t(x, u)$ is differentiable on X,

2. for all $t \in \mathbb{N}$, for all $x_t \in X$, for all $u_t^{'}, u_t^{''} \in U$, for all $\alpha \in \langle 0, 1 \rangle$, there exists $u_t \in U$ such that

$$\begin{aligned} f^0(x_t, u_t) &\geq \alpha f^0(x_t, u_t^{'}) + (1 - \alpha) f^0(x_t, u_t^{''}) \\ F_t(x_t, u_t) &= \alpha F_t(x_t, u_t^{'}) + (1 - \alpha) F_t(x_t, u_t^{''}), \end{aligned} \tag{2.5}$$

3. for any compact set $C \subset X$, there exists a constant K_C such that for all $t \in \mathbb{N}$, for all $x, x' \in C$, for all $u \in U$, $\|F_t(x, u)\| \leq K_C$ and $\|D_{x_t} F_t(x, u) - D_{x_t} F_t(x', u)\| \leq K_C \|x - x'\|$,

4. there exists $r > 0$ such that $B(\hat{x}, r) \subset X'$, where X' is the set of the bounded sequences which are in the interior of X, and for all $(x_t, u_t) \in B(\hat{x}, r) \times U$:

$$\sup_{t \geq 0} \| D_{x_t} F_t(x_t, u_t) \| < 1. \tag{2.6}$$

We call conditions (2.5) the Ioffe and Tihomirov condition [17] and they generalize the usual convexity condition used to guarantee a strong Pontryagin maximum principle (see Theorem 1.2). Blot and Hayek also consider an autonomous problem, which requires weaker conditions, namely

1. for all $u \in U$, the mapping $x \to f^0(x, u)$ is of class C^1 on X,
2. for all $t \in \mathbb{N}$, $\forall x_t \in X$, $\forall u_t^{'}, u_t^{''} \in U$, $\forall \alpha \in \langle 0, 1 \rangle$, there exists $u_t \in U$ such that

$$\begin{aligned} f^0(x_t, u_t) &\geq \alpha f^0(x_t, u_t^{'}) + (1-\alpha) f^0(x_t, u_t^{''}) \\ F(x_t, u_t) &= \alpha F(x_t, u_t^{'}) + (1-\alpha) F(x_t, u_t^{''}), \end{aligned}$$

3.

$$\sup_{t \geq 0} \| D_{x_t} F(\hat{x}_t, \hat{u}_t) \| < 1. \tag{2.7}$$

So, the regularity condition from Blot and Chebbi [5] is replaced by the supremum condition (2.6) or (2.7).

In both papers, the authors choose the approach of establishing Pontryagin maximum principle directly. As we have shown above, it requires extra conditions replacing the usual convexity conditions from Theorem 1.2. However, our motivation comes from economic problems that frequently do not satisfy them. Therefore we focus our attention on deriving pseudo-Pontryagin maximum principle in the spirit of Beran [7]. He is motivated by Blot and Hayek [6], but he considers the space ℓ_1 instead of ℓ_∞. Although, this space is not suitable for the problems with discount which we study, his work provides significant results useful for our research.

He considers the problem with both linear autonomous and general dynamics and later he extends the results to the problem with constraint on $\mathbf{u}$. Hence in comparison to the problem (2.4), control variable $u_t \in \operatorname{int} U$, there is no discount and $(\mathbf{x}, \mathbf{u}) \in \ell_1^n \times \ell_1^m$. The problem is as follows

$$\begin{aligned} J(\mathbf{x}, \mathbf{u}) &= \sum_{t=0}^{\infty} f^0(x_t, u_t) \to \max \\ x_{t+1} &= F_t(x_t, u_t), \quad t \in \mathbb{N}_0 \\ x_0 &= \bar{x}. \end{aligned}$$

He also applies the closed range theorem (Theorem 1.3), but in the form 2 $\implies$ 4. He imposes the condition that $\mathcal{L}^* = (\boldsymbol{\sigma} - \mathbf{A}^*, -\mathbf{B}^*)$ has closed range and formulate the necessary conditions of optimality in the form

$$\begin{aligned} \exists \psi = \{\psi_t\}_{t\in\mathbb{N}} \in (\ell_1^n)^* = \ell_\infty^n : D_{x_t} f(\hat{x}_t, \hat{u}_t) &= \psi_{t-1} - A_t{}^* \psi_t \quad \forall t \in \mathbb{N}_0 \\ D_{u_t} f(\hat{x}_t, \hat{u}_t) &= -B_t{}^* \psi_t \quad \forall t \in \mathbb{N}_0, \end{aligned}$$

where $A_t = D_{x_t} F_t(x_t, u_t)$ and $B_t = D_{u_t} F_t(x_t, u_t)$ and $\mathbf{A} = (A_0, A_1, \dots)$, $\mathbf{B} = (B_0, B_1, \dots)$ and $(\boldsymbol{\sigma}\mathbf{x})_t = x_{t+1}$ for any $t \in \mathbb{N}_0$. However, the research on when the condition is satisfied is incomplete.

2.2 Dual Space of ℓ_∞

Following the approach of Beran [7], we can obtain a functional $\boldsymbol{\psi}$ satisfying

$$\exists \boldsymbol{\psi} \in (\ell_\infty^n)^* : DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = (\pi_0, \boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})^* \boldsymbol{\psi}, \tag{2.8}$$

where $\pi_0(\mathbf{x}, \mathbf{u}) = x_0$.

However, $\boldsymbol{\psi}$ may not be represented by a sequence, as $(l_\infty^n)^* = \ell_1^n \oplus \ell_s^n$ and elements of ℓ_s are bounded additive scalar-valued measures on $\mathbb{N}$.

Remark 1. The elements of ℓ_s can be characterized as following: let $\mathbf{e}$ be the unit element in ℓ_∞, i.e. $e_t = 1$ for all t, and let

$$G = \{\boldsymbol{\psi} \in (\ell_\infty)^*, \boldsymbol{\psi}\mathbf{e} = 1\}$$
$$M = \{\boldsymbol{\psi} \in G, \boldsymbol{\psi}\mathbf{e}^A = 0 \text{ or } 1, \text{ for all } A \subset \mathbb{N}_0\},$$

where $\mathbf{e}^A$ with elements

$$e_t^A = \begin{cases} 1, & \text{if } t \in A \\ 0, & \text{if } t \notin A, \end{cases}$$

is called the indicator function of A. G is convex and compact in weak* topology on ℓ_∞^* induced by ℓ_∞ and M is the set of extreme points of G, i.e. its elements cannot be expressed as a proper convex combinations of other distinct elements of G. Therefore it can be shown that G is the weak* convex hull of M. Thus, the elements of ℓ_s are of the form

$$\sum_{t=0}^{n} \alpha_t \psi_t \quad \alpha_t \in \mathbb{R}, \psi \in M$$

and weak* limits of such sums. For the variety of examples of ℓ_s elements see e.g. Yosida and Hewitt [12].

Dechert in [21] describes an interesting property of ℓ_s For $\boldsymbol{\psi} \in M$, let

$$N_{\boldsymbol{\psi}} = \{A \subset \mathbb{N}_0, \boldsymbol{\psi}\mathbf{e}^A = 1\}.$$

Hence

$$\boldsymbol{\psi}(\mathbf{x}\mathbf{e}^A) = \boldsymbol{\psi}\mathbf{x}(\boldsymbol{\psi}\mathbf{e}^A) = \boldsymbol{\psi}x \quad \mathbf{x} \in \ell_\infty, A \subset N_{\boldsymbol{\psi}},$$

i.e. if for some $A \subset N_{\boldsymbol{\psi}}$ the sequence $\mathbf{x} = \{x_t\}_{t \in A}$ converges to $\mathbf{x}_0$, then $\boldsymbol{\psi}\mathbf{x} = \mathbf{x}_0$. As a special case, if $A \subset \mathbb{N}_0$ is a finite subset then $A^C \in N_{\boldsymbol{\psi}}$ for any $\boldsymbol{\psi} \in M$ and so $(\mathbf{y}\mathbf{e}^A) = 0$, for all $\mathbf{y} \in \ell_\infty$, hence

$$\boldsymbol{\psi}(\mathbf{x}+\mathbf{y}\mathbf{e}^A) = \boldsymbol{\psi}\mathbf{x} \quad \text{for all } \mathbf{x} \in \ell_\infty.$$

This property is used in methods by Dechert [21], Blot and Hayek [6] and Le Van, Saglam [18].

Blot and Hayek [6] applied the following lemma.

Lemma 2.1. *If $\boldsymbol{\psi}^s \in \ell_s$, then there exists $k \in \mathbb{R}$ such that for all $\mathbf{x} \in c$, $\boldsymbol{\psi}^s\mathbf{x} = k\lim_{t\to\infty} x_t$.*

Proof. The proof can be found in [23]. □

We adapt this approach, therefore we describe the method by Blot and Hayek [6] more thouroughly.

Firstly, they show that under the assumptions (2.5), (2.6), the mapping $\mathbf{x} \to J(\mathbf{x},\mathbf{u})$ is of class C^1 for all $\mathbf{u}$, for all $\mathbf{x}^\infty = \{x_t^\infty\}_{t\in\mathbb{N}} \in \ell_\infty^n$, $D_\mathbf{x}J(\mathbf{x},\mathbf{u})\mathbf{x}^\infty = \sum_{t=0}^\infty \delta^t D_{x_t} f^0(x_t,u_t)x_t^\infty$. They set $\mathbf{F}(\mathbf{x},\mathbf{u}) = \{F_t(x_t,u_t) - x_{t+1}\}_{t\in\mathbb{N}_0}$, showing that $\mathbf{F}(\mathbf{x},\mathbf{u}) \in \ell_\infty^n$. The supremum condition (2.6) ensures that the mapping $\mathbf{x} \to \mathbf{F}(\mathbf{x},\mathbf{u})$ is of class C^1 for all $\mathbf{u}$ and for any $\mathbf{x}^\infty \in \ell_\infty^n$ $D_\mathbf{x}\mathbf{F}(\mathbf{x},\mathbf{u}) = \{D_{x_t}F_t(x_t,u_t)x_t^\infty - x_{t+1}^\infty\}_{t\in\mathbb{N}_0}$. They also show that $J, \mathbf{F}$ fulfill the Ioffe-Tihomirov condition (2.5).

Then they establish adjoint equation and strong Pontryagin maximum principle with $\boldsymbol{\psi} \in (\ell_\infty)^*$ in the following formulation
there exists $\psi_0 \in \mathbb{R}, \psi_0 \geq 0, \boldsymbol{\psi} \in (\ell_\infty)^*$, not all zero, such that:

$$\psi_0 D_\mathbf{x}J(\hat{\mathbf{x}},\hat{\mathbf{u}}) + D_\mathbf{x}\mathbf{F}(\hat{\mathbf{x}},\hat{\mathbf{u}})\boldsymbol{\psi} = 0 \tag{2.9}$$

$$(\psi_0 J + \langle\boldsymbol{\psi},\mathbf{F}\rangle)(\hat{\mathbf{x}},\hat{\mathbf{u}}) \geq (\psi_0 J + \langle\boldsymbol{\psi},\mathbf{F}\rangle)(\hat{\mathbf{x}},\mathbf{u}), \quad \text{for all } \mathbf{u} \in U. \tag{2.10}$$

Then they split the adjoint variable into two parts $\boldsymbol{\psi} = \boldsymbol{\psi}^1 + \boldsymbol{\psi}^s$, where $\boldsymbol{\psi}^1 \in \ell_1$ and

$\boldsymbol{\psi}^s \in \ell_s$. So (2.9) can be rewritten to

$$\sum_{t=0}^{\infty}\psi_0\delta^t D_{x_t}f^0(\hat{x}_t,\hat{u}_t)x_t^\infty + \sum_{t=0}^{\infty}\langle\psi_{t+1}^1, D_{x_t}F_t(\hat{x}_t,\hat{u}_t)x_t^\infty\rangle - \sum_{t=0}^{\infty}\langle\psi_{t+1}^1, x_{t+1}^\infty\rangle$$
$$+ \langle\boldsymbol{\psi}_s, \{D_{x_t}F_t(\hat{x}_t,\hat{u}_t)x_t^\infty - x_{t+1}^\infty\}_{t\in\mathbb{N}_0}\rangle = 0, \quad \text{for all } \mathbf{x}^\infty \in \ell_\infty^n \text{ with } x_0^\infty = 0.$$

And they obtain

$$\begin{aligned}\sum_{t=0}^{\infty}\langle\psi_0\delta^t D_{x_t}f^0(\hat{x}_t,\hat{u}_t) + D_{x_t}F_t(\hat{x}_t,\hat{u}_t)\psi_{t+1}^1 - \psi_t^1, x_t^\infty\rangle = \\ -\langle\boldsymbol{\psi}^s, \{D_{x_t}F_t(\hat{x}_t,\hat{u}_t)x_t^\infty - x_{t+1}^\infty\}_{t\geq 0}\rangle, \quad \mathbf{x}^\infty \in \ell_\infty \text{ with } x_0^\infty = 0.\end{aligned} \tag{2.11}$$

Let z be chosen arbitrarily in $\mathbb{R}^n$ and let consider the sequence $\mathbf{x}^\infty \in \ell_\infty$ defined as follows

$$x_s^\infty = \begin{cases} z, & \text{if } s = t \\ 0, & \text{if } s \neq t. \end{cases}$$

So one has $D_{x_s}F_s(\hat{x}_s,\hat{u}_s)x_s^\infty - x_{s+1}^\infty = 0$ if $s \geq t+1$, hence $\{D_{x_s}F_s(\hat{x}_s,\hat{u}_s)x_s^\infty - x_{s+1}^\infty\}_{s\in\mathbb{N}_0} \in c_0 \subset c$.

Thus, according to the Lemma 2.1

$$\langle\boldsymbol{\psi}^s, \{D_{x_s}F_s(\hat{x}_s,\hat{u}_s)x_s^\infty - x_{s+1}^\infty\}_{s\geq 0}\rangle = \langle k, \lim_{s\to\infty}(D_{x_s}F_s(\hat{x}_s,\hat{u}_s)x_s^\infty - x_{s+1}^\infty)\rangle = 0.$$

Therefore we have for all $z \in \mathbb{R}^n$

$$\langle\psi_0\delta^t D_{x_t}f^0(\hat{x}_t,\hat{u}_t) + D_{x_t}F_t(\hat{x}_t,\hat{u}_t)\psi_{t+1}^1 - \psi_t^1, z\rangle = 0,$$

which implies $\psi_t^1 = \psi_0\delta^t D_{x_t}f^0(\hat{x}_t,\hat{u}_t) + D_{x_t}F_t(\hat{x}_t,\hat{u}_t)\psi_{t+1}^1$, for all $t \geq 1$.

Chapter 3

F-differentiability of the objective function

In this chapter, we show that the objective function $J(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{\infty} \delta^t f^0(x_t, u_t)$ is Fréchet differentiable. At first, we define the differentiability and lemmas, then we prove the proposition.

Definition 3.1. Let $J : U \to Y$, where X, Y are Banach spaces and $U \subset X$ is open. Let $x \in U$ and $h \in X$. The *directional derivative* at x of a function $J(x)$ along vector h is defined by

$$\partial_h J(x) = \lim_{\tau \to 0} \frac{1}{\tau}[J(x + \tau h) - J(x)],$$

if it exists.

Definition 3.2. Let $J : U \to Y$, where X, Y are Banach spaces and $U \subset X$ is open, let $x \in U$ and let $\partial_h J(x)$ exists for all h. We call the function J *Gâteaux differentiable* in x if the map $h \to \partial_h J(x)$ is linear and bounded. It is defined by

$$dJ(x)h = \partial_h J(x).$$

Definition 3.3. Let $J : U \to Y$, where X, Y are Banach spaces and $U \subset X$ is open, let $x \in U$. We call the function J *Fréchet differentiable* in x if there exists a linear bounded operator $DJ(x)$ such that

$$\lim_{\|h\|\to 0} \frac{1}{\|h\|}[J(x+h) - J(x) - DJ(x)h] = 0.$$

Lemma 3.1. ***(Hadamard's lemma)***
Let X, Y be Banach spaces and $f : U \to Y$ be a Gâteaux differentiable mapping. If $x + \eta h \in U$ for $\eta \in \langle 0, 1\rangle$, then one has

$$f(x+h) - f(x) = \int_0^1 df(x+\eta h)h d\eta = \left[\int_0^1 df(x+\eta h)d\eta\right] h,$$

where the integral is Riemann.

Proof. Let denote $F(\eta) = f(x+\eta h)$, where $\eta \in \langle 0, 1\rangle$. Then $F(1)-F(0) = f(x+h)-f(x)$ and F is differentiable, so

$$f(x+h) - f(x) = F(1) - F(0) = \int_0^1 F'(\eta)d\eta.$$

As f is Gâteaux differentiable mapping we have

$$\int_0^1 F'(\eta)d\eta = \int_0^1 \partial_h f(x+\eta h)d\eta = \int_0^1 df(x+\eta h)hd\eta.$$

□

Lemma 3.2. *Let $J : U \to Y$, where X, Y are Banach spaces and $U \subset X$ is open, let $x \in U$. If J is Gâteaux differentiable and the Gâteaux derivative is continuous on a neighborhood V of x, then J is Fréchet differentiable at x.*

Proof. Since dJ is continuous, for a given ε, there exists $\delta > 0$ such that if $k \in X$, $\|k\| \leq \delta$, then $J(x+k) \in U$ and

$$\|\partial_h J(x+k) - \partial_h J(x)\| = \|[dJ(x+k) - dJ(x)]h\| \leq \varepsilon \|h\|,$$

for all $h \in X$. Let $\|h\| \leq \delta$, then according to Hadamard's lemma we have

$$\|J(x+h) - J(x) - dJ(x)h\| = \left\| \int_0^1 dJ(x+\eta h) d\eta h - dJ(x)h \right\|$$
$$= \left\| \int_0^1 [dJ(x+\eta h) - dJ(x)] h d\eta \right\| \leq \left\| \int_0^1 [dJ(x+\eta h) - dJ(x)] d\eta \right\| \|h\| \leq \varepsilon \|h\|.$$

□

Proposition 3.1. The function $J : \ell_p^n \times \ell_p^n \to \mathbb{R}$, $p \in \langle 1, \infty \rangle$, defined by $J(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{\infty} \delta^t f^0(x_t, u_t)$, where $x_t \in X \subset \mathbb{R}^n$, $u_t \in U \subset R^m$ and $f^0 \in C^1(X \times U, \mathbb{R})$ is Fréchet differentiable.

Proof. We carry out the proof in three steps. Firstly, we show that there exists $\partial_{\mathbf{h}} J(\mathbf{x}, \mathbf{u})$ for all $\mathbf{h} \in \ell_p^{n+m}$. Then we prove that the map $\mathbf{h} \to \partial_{\mathbf{h}} J(\mathbf{z})$ is linear and bounded, thus it is Gâteaux differentiable. Finally, we show that it is Fréchet differentiable.

Let us simplify $(\mathbf{x}, \mathbf{u}) = \mathbf{z}$ and $(x_t, u_t) = z_t$, $\mathbf{h} = \{h_t\}_{t \in \mathbb{N}_0}$. Since $\mathbf{x} \in \ell_p^n$ and $\mathbf{u} \in \ell_p^m$, the conditions

$$\sum_{t=0}^{\infty} |x_t|^p < \infty, \sum_{t=0}^{\infty} |u_t|^p < \infty \text{ for } p \in \langle 1, \infty)$$

or

$$\sup_{t \in \mathbb{N}_0} |x_t| < \infty, \sup_{t \in \mathbb{N}_0} |u_t| < \infty \text{ for } p = \infty$$

are fulfilled. So $|x_t| < \infty$, $|u_t| < \infty$ for all t. Hence there exist compact sets X_0 and U_0 such that $x_t \in X_0, u_t \in U_0$. As a continuous function on compact set is bounded, we have

$$|D_{x_t} f^0| < C$$
$$|D_{u_t} f^0| < C \quad \text{on } X_0 \times U_0, \text{ for some } C > 0,$$

so

$$|D_{z_t}f^0| \le |D_{x_t}f^0| + |D_{u_t}f^0| < 2C. \tag{3.1}$$

And using the same argumentation $D_{x_t}f^0$, $D_{u_t}f^0$ and so $D_{z_t}f^0$ are also uniformly continuous.

1. *We will show that there exists directional derivative of $J(\mathbf{z})$ along any $\mathbf{h} \in \ell_p^{n+m}$.*

$$\begin{aligned}
\partial_{\mathbf{h}} J(\mathbf{z}) &= \lim_{\tau\to 0} \frac{1}{\tau}[J(\mathbf{z}+\tau\mathbf{h}) - J(\mathbf{z})] \\
&= \lim_{\tau\to 0} \frac{1}{\tau}\left[\sum_{t=0}^{\infty} \delta^t f^0(z_t+\tau h_t) - \sum_{t=0}^{\infty} \delta^t f^0(z_t)\right] \\
&= \lim_{\tau\to 0} \frac{1}{\tau}\sum_{t=0}^{\infty} \delta^t \left[f^0(z_t+\tau h_t) - f^0(z_t)\right].
\end{aligned}$$

We want to interchange the summation and limit, so that we obtain $\partial_{\mathbf{h}} J(\mathbf{z}) = \sum_{t=0}^{\infty} \partial_{h_t} f(z_t)$. Therefore we prove absolute convergence of the series. For τ sufficiently small one has $(z_t + \eta\tau h_t) \in X_0 \times U_0$. Therefore we can employ Lemma 3.1 (Hadamard's lemma) and in case $p = 1$ we obtain

$$\begin{aligned}
\sum_{t=0}^{\infty} |\delta^t[f^0(z_t+\tau h_t) - f^0(z_t)]| &= \sum_{t=0}^{\infty} |\delta^t(\tau h_t)\int_0^1 D_{z_t} f^0(z_t+\eta(\tau h_t))d\eta| \\
&= \sum_{t=0}^{\infty} \delta^t |\tau h_t| . \left|\int_0^1 D_{z_t} f^0(z_t+\eta\tau h_t)d\eta\right| \\
&\le 2C\sum_{t=0}^{\infty} \delta^t|\tau h_t| \left|\int_0^1 d\eta\right| = 2C\sum_{t=0}^{\infty}\delta^t|\tau||h_t| \\
&= 2C|\tau|\sum_{t=0}^{\infty}\delta^t|h_t| \le \frac{2C|\tau|}{1-\delta}\sum_{t=0}^{\infty}|h_t| = \frac{2C|\tau|}{1-\delta}\|\mathbf{h}\|_1 < \infty.
\end{aligned}$$

In case $p \in (1, \infty\rangle$ we proceed similarly and we employ Hölder inequality assuming that $\frac{1}{p} + \frac{1}{q} = 1$ and for $p = \infty$ we define $\frac{1}{p} = 0$.

$$\begin{aligned}\sum_{t=0}^{\infty}\left|\delta^t[f^0(z_t+\tau h_t)-f^0(z_t)]\right| &= \sum_{t=0}^{\infty}\left|\delta^t\tau h_t\int_0^1 D_{z_t}f^0(z_t+\eta(\tau h_t))d\eta\right| \\ &\le \|\mathbf{h}\|_p\left(\sum_{t=0}^{\infty}\left|\delta^t\tau\int_0^1 D_{z_t}f^0(z_t+\eta(\tau h_t))d\eta\right|^q\right)^{\frac{1}{q}} \\ &\le \|\mathbf{h}\|_p\left(\sum_{t=0}^{\infty}\left|\delta^t\tau\int_0^1 2Cd\eta\right|^q\right)^{\frac{1}{q}} = \|\mathbf{h}\|_p 2C|\tau|\left(\sum_{t=0}^{\infty}\delta^{tq}\right)^{\frac{1}{q}} \\ &= 2C|\tau|\left(\frac{1}{1-\delta^q}\right)^{\frac{1}{q}}\|\mathbf{h}\|_p<\infty\end{aligned}$$

Now, we can interchange limit and summation and we obtain

$$\partial_{\mathbf{h}}J(\mathbf{z})=\sum_{t=0}^{\infty}\partial_{h_t}\delta^t f^0(z_t)=\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)h_t.$$

Thus $\partial_{\mathbf{h}}J(\mathbf{z})$ exists.

2. *Now, we prove that the map $\mathbf{h}\to\partial_{\mathbf{h}}J(\mathbf{z})$ is linear and bounded.*

To check linearity we write

$$\begin{aligned}\partial_{\alpha\mathbf{h}+\beta\mathbf{g}}J(\mathbf{z}) &= \sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)(\alpha h_t+\beta g_t)=\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)(\alpha h_t)+\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)(\beta g_t) \\ &= \alpha\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)h_t+\beta\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)g_t=\alpha\partial_{\mathbf{h}}J(\mathbf{z})+\beta\partial_{\mathbf{g}}J(\mathbf{z})\end{aligned}$$

Next, we prove boundness in case $p=1$

$$\begin{aligned}|\partial_{\mathbf{h}}J(\mathbf{z})| &\le \left|\sum_{t=0}^{\infty}\delta^t D_{z_t}f^0(z_t)h_t\right| \le \sum_{t=0}^{\infty}\delta^t\left|D_{z_t}f^0(z_t)h_t\right| \\ &\le \frac{1}{1-\delta}\sum_{t=0}^{\infty}|D_{z_t}f^0(z_t)||h_t| \le \frac{2C}{1-\delta}\|\mathbf{h}\|_1<\infty.\end{aligned}$$

And finally we show that if $p\in(1,\infty\rangle$ the map is bounded as well. Again, we apply

Hölder inequality.

$$|\partial_{\mathbf{h}} J(\mathbf{z})| \leq \left| \sum_{t=0}^{\infty} \delta^t D_{z_t} f^0(z_t) h_t \right| \leq \sum_{t=0}^{\infty} \left| \delta^t D_{z_t} f^0(z_t) h_t \right| \leq \|\mathbf{h}\|_p \left(\sum_{t=0}^{\infty} |\delta^t D_{z_t} f^0(z_t)|^q \right)^{\frac{1}{q}}$$
$$\leq \|\mathbf{h}\|_p 2C \left(\sum_{t=0}^{\infty} \delta^t \right)^{\frac{1}{q}} = \|\mathbf{h}\|_p 2C \left(\frac{1}{1-\delta^q} \right)^{\frac{1}{q}} < \infty.$$

So, J is Gâteaux differentiable for any $p \in \langle 1, \infty \rangle$.

3. *To show that J is also Fréchet differentiable we employ Lemma 3.2.*

We have to prove that the Gâteaux derivative is continuous, i.e. for a given $\varepsilon > 0$, there exists $\delta > 0$, such that for all $\mathbf{y} \in \ell_p^{n+m}$, if $(\mathbf{z} - \mathbf{y}) \in X$, $\|\mathbf{z} - \mathbf{y}\|_p < \delta$ then

$$|\partial_{\mathbf{h}} J(\mathbf{z}) - \partial_{\mathbf{h}} J(\mathbf{y})| \leq \varepsilon \|\mathbf{h}\|_p, \tag{3.2}$$

for all $\mathbf{h} \in \ell_p^{n+m}$. For $p \in \langle 1, \infty)$ we have

$$|z_t - y_t| \leq \left(\sum_{t=0}^{\infty} |z_t - y_t|^p \right)^{1/p} = \|\mathbf{z} - \mathbf{y}\|_p < \lambda$$

and for $p = \infty$

$$|z_t - y_t| \leq \sup_{t \in \mathbb{N}_0} |z_t - y_t| = \|\mathbf{z} - \mathbf{y}\|_\infty < \lambda.$$

Thus, for any p we have $|z_t - y_t| < \lambda$.

We have already shown that $D_{z_t} f$ is uniformly continuous on $X_0 \times U_0$, so for a given $\varepsilon > 0$ there exists $\lambda > 0$ such that if $|z_t - y_t| < \lambda$ then

$$|D_{z_t} f(z_t) - D_{y_t} f(y_t)| \leq \varepsilon^* \quad \text{on } X_0 \times U_0.$$

We can rewrite the left side of inequality (3.2) to

$$
\begin{aligned}
|\partial_{\mathbf{h}} J(\mathbf{z}) - \partial_{\mathbf{h}} J(\mathbf{y})| &= \left| \sum_{t=0}^{\infty} \delta^t D_{z_t} f^0(z_t) h_t - \sum_{t=0}^{\infty} \delta^t D_{y_t} f^0(y_t) h_t \right| \\
&= \left| \sum_{t=0}^{\infty} \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] h_t \right|
\end{aligned}
$$

Next, for $p = 1$ we employ the uniform continuity and write

$$
\begin{aligned}
\left| \sum_{t=0}^{\infty} \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] h_t \right| &\leq \sum_{t=0}^{\infty} \left| \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] h_t \right| \\
&\leq \frac{1}{1-\delta} \sum_{t=0}^{\infty} \left| [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] \right| |h_t| \\
&\leq \frac{\varepsilon^*}{1-\delta} \sum_{t=0}^{\infty} |h_t| = \frac{\varepsilon^*}{1-\delta} \|\mathbf{h}\|_1.
\end{aligned}
$$

Applying Hölder inequality we have for $p \in (1, \infty\rangle$

$$
\begin{aligned}
\left| \sum_{t=0}^{\infty} \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] h_t \right| &\leq \sum_{t=0}^{\infty} \left| \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] h_t \right| \\
&\leq \|\mathbf{h}\|_p \left(\sum_{t=0}^{\infty} \left| \delta^t [D_{z_t} f^0(z_t) - D_{y_t} f^0(y_t)] \right|^q \right)^{\frac{1}{q}} \\
&\leq \varepsilon^* \|\mathbf{h}\|_p \left(\sum_{t=0}^{\infty} \delta^{tq} \right)^{\frac{1}{q}} = \varepsilon^* \|\mathbf{h}\|_p \left(\frac{1}{1-\delta^q} \right)^{\frac{1}{q}}
\end{aligned}
$$

By setting $\varepsilon = \varepsilon^* \left(\frac{1}{1-\delta^q} \right)^{\frac{1}{q}}$ we obtain the required inequality. □

Chapter 4

The Method

In this chapter we desribe the method establishing the pseudo-Pontryagin maximum principle for discrete-time optimal control problems on infinite horizon. At first we consider the problem with linear autonomous dynamics, then we extend our results to general dynamics.

4.1 The Linear Autonomous Problem

We describe our method on the infinite-horizon discrete-time optimal control model with linear autonomous dynamics

$$J(\mathbf{x},\mathbf{u}) = \sum_{t=0}^{\infty} \delta^t f^0(x_t, u_t) \to \max \tag{4.1}$$

$$x_{t+1} = Ax_t + Bu_t + d \quad \text{for all } t \in \mathbb{N}_0 \tag{4.2}$$

$$x_0 = \bar{x}, \tag{4.3}$$

where $\bar{x}$,$d \in \mathbb{R}^n$, $n \times n$ matrix A, $n \times m$ matrix B and discount $\delta \in (0,1)$ are given. We denote $x_t \in \mathbb{R}^n = X$, $\mathbf{x} = \{x_t\}_{t=0}^{\infty}$, $u_t \in \mathbb{R}^m = U$ $\mathbf{u} = \{u_t\}_{t=0}^{\infty}$, objective function $f^0 \in C^1(X \times U, \mathbb{R})$.

Firstly, we construct perturbations along the optimal solution, then we formulate necessary conditions of optimality.

Let $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ be optimal solution of problem (4.1) - (4.3). A pair $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \ell_\infty^{n+m}$ is called admissible, if for all $\varepsilon > 0$ it holds

$$\begin{aligned}\hat{x}_0 + \varepsilon\alpha_0 &= \bar{x} \\ \hat{x}_{t+1} + \varepsilon\alpha_{t+1} &= A(\hat{x}_t + \varepsilon\alpha_t) + B(\hat{u}_t + \varepsilon\beta_t) + d \quad \text{for all } t \in \mathbb{N}_0,\end{aligned}$$

i. e. $\{\hat{x}_t + \varepsilon\alpha_t, \hat{u}_t + \varepsilon\beta_t\}$ satisfies (4.2) and (4.3).

Because of (4.3), one has $\alpha_0 = 0$. Next, we apply equation $\hat{x}_{t+1} = A\hat{x}_t + B\hat{u}_t + d$ and the system can be rewritten to

$$\begin{aligned}\alpha_0 &= 0 \\ \alpha_{t+1} &= A\alpha_t + B\beta_t \quad \text{for all } t \in \mathbb{N}_0.\end{aligned}$$

From the definition of an admissible vector, $J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})$ cannot increase with ε (≥ 0) from the maximum. We have already shown that J is Fréchet differentiable in ℓ_∞^{n+m} (Proposition 3.1), therefore

$$\begin{aligned}\frac{\partial}{\partial\varepsilon} J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} \leq 0 \\ \frac{\partial}{\partial\varepsilon} J(\hat{\mathbf{x}} - \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} - \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} \leq 0\end{aligned} \tag{4.4}$$

As $\frac{\partial}{\partial\varepsilon} J(\hat{\mathbf{x}} - \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} - \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} = -\frac{\partial}{\partial\varepsilon} J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})|_{\varepsilon=0}$, (4.4) can be rewritten to

$$\begin{aligned}0 = \frac{\partial}{\partial\varepsilon} J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} &= \sum_{t=0}^{\infty} \delta^t [D_{x_t} f^0(\hat{x}_t, \hat{u}_t)\alpha_t + D_{u_t} f^0(\hat{x}_t, \hat{u}_t)\beta_t] \\ &= DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal.\end{aligned}$$

This notation can be simplified by defining $\mathbf{A}$, $\mathbf{B}$ and by introducing a vector of shift operators $\boldsymbol{\sigma}$, such that $(\mathbf{A}\boldsymbol{\alpha})_t = A\alpha_t$, $(\mathbf{B}\boldsymbol{\beta})_t = B\beta_t$ and $(\boldsymbol{\sigma}\alpha)_t = \alpha_{t+1}$ and we obtain

$$\alpha_0 = 0 \tag{4.5}$$

$$(\boldsymbol{\sigma} - \mathbf{A})\boldsymbol{\alpha} - \mathbf{B}\boldsymbol{\beta} = (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0. \tag{4.6}$$

Remark 2. Let us define an operator $\pi_0 = (\pi_0^x, \mathbf{0})$ such that $\pi_0(\mathbf{x}, \mathbf{u})^\intercal = x_0$ and an operator $\mathcal{L} : \ell_\infty^n \times \ell_\infty^m \to \ell_\infty^n$, $\mathcal{L} = (\pi_0, (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B}))$. Then conditions (4.5) and (4.6) can be replaced by $\mathcal{L}(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0$ or $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{N}(\mathcal{L})$.

In order to apply closed range theorem, $\mathcal{L}$ needs to be bounded.

Proposition 4.1. Let $\mathbf{A}$ and $\mathbf{B}$ be general linear operators. Then $\mathcal{L} : \ell_p^n \times \ell_p^m \to \ell_p^n$, $\mathcal{L} = (\pi_0, (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B}))$ is bounded linear operator for any $p \in \langle 0, \infty \rangle$.

Proof. The proof of linearity is trivial. We prove that $\boldsymbol{\sigma}, \mathbf{A}, \mathbf{B}$ and π_0 are bounded, hence $\mathcal{L}$ is bounded.

$$\|\pi_0\| = \sup_{\|\mathbf{x}\|_p = 1} \|\pi_0 \mathbf{x}\|_p = \sup_{\|\mathbf{x}\|_p = 1} |x_0| \le 1$$

$$\|\boldsymbol{\sigma}\|_p = \sup_{\|\mathbf{x}\|_p = 1} \|\boldsymbol{\sigma}\mathbf{x}\|_p = \sup_{\|\mathbf{x}\|_p = 1} \|(x_1, x_2, \dots)\|_p \le \sup_{\|\mathbf{x}\|_p = 1} \|x\|_p = 1$$

As $\mathbf{A} = (A_1, A_2, \dots)$, where A_t are $n \times n$ matrix, $\|A_t\| \le M < \infty$ for all $t \in \mathbb{N}$

$$\begin{aligned}\|\mathbf{A}\|_p = \sup_{\|x\| = 1} \|\mathbf{A}\mathbf{x}\|_p &= \sup_{\|\mathbf{x}\|_p = 1} \left(\sum_{t=0}^{\infty} |A_t x_t|^p \right)^{\frac{1}{p}} \\ &\le \sup_{\|\mathbf{x}\|_p = 1} \left(\sum_{t=0}^{\infty} |A_t|^p |x_t|^p \right)^{\frac{1}{p}} \le M \sup_{\|x\| = 1} \|\mathbf{x}\|_p = M,\end{aligned}$$

in case $p \neq \infty$ and if $p = \infty$

$$\|\mathbf{A}\|_\infty = \sup_{\|\mathbf{x}\|_\infty = 1} \sup_{t \ge 0} |A_t x_t| \le M \sup_{\|\mathbf{x}\|_\infty = 1} \|\mathbf{x}\|_\infty = M$$

And by the same argumentation $\mathbf{B}$ is bounded. $\square$

Theorem 4.1. *(Necessary conditions of optimality)*

Assume that the operator $\mathcal{L} = (\pi_0, \boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})$ *has closed range. Then* $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0$ *for all admissible* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ *if only if there exists* $\boldsymbol{\phi} \in (\ell_\infty)^*$ *such that*

$$DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \mathcal{L}^* \boldsymbol{\phi} \tag{4.7}$$

Moreover, if one has $\boldsymbol{\phi} = \boldsymbol{\psi} + \boldsymbol{\phi}^s$*, where* $\boldsymbol{\psi} = \{\psi_t\}_{t \in \mathbb{N}_0} \in \ell_1$ *and* $\boldsymbol{\phi}^s \in \ell_s$*, then* $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0$ *for all admissible* $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ *if*

$$\begin{aligned} D_{x_t} f^0(\hat{x}_t, \hat{u}_t) &= \psi_{t-1} - \delta A^* \psi_t \text{ for all } t \in \mathbb{N} \\ D_{u_t} f^0(\hat{x}_t, \hat{u}_t) &= -\delta B^* \psi_t \text{ for all } t \in \mathbb{N}_0. \end{aligned} \tag{4.8}$$

Proof. At first, we show that there exists $\boldsymbol{\psi} \in (\ell_\infty)^*$ such that equation (4.7) holds, then we show that in terms it can be rewritten to the system (4.8).

According to the Proposition 4.1 $\mathcal{L} = (\pi_0, (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B}))$ is bounded and by the assumption it has closed range. Hence according to the closed range theorem (Theorem 1.3)

$$\mathcal{R}(\mathcal{L}^*) = \{x^* \in X^* : \langle x^*, (\boldsymbol{\alpha}, \boldsymbol{\beta}) \rangle = 0 \text{ for all } (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{N}(\mathcal{L})\},$$

so $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \in \mathcal{R}(\mathcal{L}^*)$, i.e. there exists $\boldsymbol{\phi} \in (\ell_\infty)^*$ such that $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \mathcal{L}^* \phi$. Hence

$$\begin{aligned} D_{\mathbf{x}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) - (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})^* \boldsymbol{\phi} &= 0 \\ D_{\mathbf{u}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) + (\mathbf{0}, \mathbf{B})^* \boldsymbol{\phi} &= 0. \end{aligned}$$

So the first part of the theorem is proved.

For any sequences $\mathbf{x} \in \ell_\infty^n$ and $\mathbf{u} \in \ell_\infty^m$

$$\begin{aligned} D_{\mathbf{x}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \mathbf{x} - \langle (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})^* \boldsymbol{\phi}, \mathbf{x} \rangle &= D_{\mathbf{x}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \mathbf{x} - \langle \boldsymbol{\phi}, (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A}) \mathbf{x} \rangle = 0 \\ D_{\mathbf{u}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \mathbf{u} + \langle (\mathbf{0}, \mathbf{B})^* \boldsymbol{\phi}, \mathbf{u} \rangle &= D_{\mathbf{u}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}}) \mathbf{u} + \langle \boldsymbol{\phi}, (\mathbf{0}, \mathbf{B}) \mathbf{u} \rangle = 0. \end{aligned}$$

Now, we split $\boldsymbol{\phi} \in (\ell_\infty)^* = \ell_1 \oplus \ell_s$ to $\boldsymbol{\phi} = \boldsymbol{\phi}^{\mathbf{1}} + \boldsymbol{\phi}^s$ such that $\boldsymbol{\phi}^1 \in \ell_1$ and $\boldsymbol{\phi}^s \in \ell_s$.

$$
\begin{aligned}
D_{\mathbf{x}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}})\mathbf{x} - \langle \boldsymbol{\phi}^{\mathbf{1}}, (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})\mathbf{x} \rangle &= \langle \boldsymbol{\phi}^{\boldsymbol{s}}, (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})\mathbf{x} \rangle \\
D_{\mathbf{u}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}})\mathbf{u} + \langle \boldsymbol{\phi}^1, (\mathbf{0}, \mathbf{B})\mathbf{u} \rangle &= \langle \boldsymbol{\phi}^s, -(\mathbf{0}, \mathbf{B})\mathbf{u} \rangle.
\end{aligned}
$$

Consider the sequences $\mathbf{x}^\tau = \{x_t^\tau\}_{t\in\mathbb{N}}$ and $\mathbf{u}^\tau = \{u_t^\tau\}_{t\in\mathbb{N}}$ such that

$$
x_t^\tau = \begin{cases} z_x, & \text{if } t = \tau \\ 0, & \text{if } t \neq \tau, \tau \in \mathbb{N} \end{cases} \qquad u_t^\tau = \begin{cases} z_u, & \text{if } t = \tau \\ 0, & \text{if } t \neq \tau, \tau \in \mathbb{N}_0 \end{cases}
$$

where $z_x \in \mathbb{R}^n$ and $z_u \in \mathbb{R}^m$ are chosen arbitrary. Then

$$
\begin{aligned}
(\sigma - A){x_t}^\tau &= 0_n \\
-B{u_t}^\tau &= 0_m
\end{aligned}
$$

for all $t \geq \tau + 1$, where $0_n = (0, 0, \dots, 0) \in X$ and $0_m = (0, 0, \dots, 0) \in U$. Hence $(\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})\mathbf{x}^\tau \in c_0 \subset c$ and $-(\mathbf{0}, \mathbf{B})\mathbf{u}^\tau \in c_0 \subset c$. By Lemma 2.1 there exist $k_1, k_2 \in \mathbb{R}$ such that

$$
\begin{aligned}
\langle \boldsymbol{\phi}^{\boldsymbol{s}}, (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})\mathbf{x}^\tau \rangle &= \langle k_1, \lim_{t\to\infty} (\sigma - A){x_t}^\tau \rangle = \langle k_1, 0_n \rangle = 0 \\
\langle \boldsymbol{\phi}^s, -(\mathbf{0}, \mathbf{B})\mathbf{u}^\tau \rangle &= \langle k_2, \lim_{t\to\infty} -B{u_t}^\tau \rangle = \langle k_2, 0_m \rangle = 0.
\end{aligned}
$$

We have for all $\tau \in \mathbb{N}$ and all $z_x \in X$

$$
\begin{aligned}
0 &= D_{\mathbf{x}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}})\mathbf{x}^\tau - \langle (\pi_0^x, \boldsymbol{\sigma} - \mathbf{A})^* \boldsymbol{\phi}^{\mathbf{1}}, \mathbf{x}^\tau \rangle \\
&= \phi_0 x_0^\tau + \sum_{t=0}^{\infty} \left(\delta^\tau D_{x_\tau} f^0(\hat{x_\tau}, \hat{u_\tau}) - \phi_\tau^1 + A^* \phi_{\tau+1}^1 \right) x_t^\tau \\
&= \left(\delta^\tau D_{x_\tau} f^0(\hat{x_\tau}, \hat{u_\tau}) - \phi_\tau^1 + A^* \phi_{\tau+1}^1 \right) z_x
\end{aligned}
$$

by the same argumentation for all $\tau \in \mathbb{N}_0$ and all $z_u \in U$

$$\begin{aligned} 0 &= D_{\mathbf{u}} J(\hat{\mathbf{x}}, \hat{\mathbf{u}})\mathbf{u}^\tau + \langle \boldsymbol{\phi}^1, (\mathbf{0}, \mathbf{B})^* \mathbf{u}^\tau \rangle \\ &= \sum_{t=0}^{\infty} \left(\delta^t D_{u_t} f^0(\hat{x}_t, \hat{u}_t) + B^* \phi^1_{t+1} \right) u^\tau_t \\ &= \left(\delta^\tau D_{u_\tau} f^0(\hat{x}_\tau, \hat{u_\tau}) + B^* \phi^1_{\tau+1} \right) z_u, \end{aligned}$$

which implies that

$$\begin{aligned} \delta^t D_{x_t} f^0(\hat{x}_t, \hat{u_t}) - \phi^1_t + A^* \phi^1_{t+1} &= 0 \text{ for all } t \in \mathbb{N} \\ \delta^t D_{u_t} f^0(\hat{x}_t, \hat{u_t}) + B^* \phi^1_{t+1} &= 0 \text{ for all } t \in \mathbb{N}_0. \end{aligned}$$

Finally if we put $\psi_t = \frac{\phi^1_{t+1}}{\delta^{t+1}}$, we obtain the required equation. □

4.2 The Problem with General Dynamics

In this section, we replace the linear autonomous dynamics (4.2) by generalized dynamics $F_t \in C^1(X \times U, \mathbb{R})$ for all $t \in \mathbb{N}_0$, i.e. we consider the problem

$$J(\mathbf{x}, \mathbf{u}) = \sum_{t=0}^{\infty} \delta^t f^0(x_t, u_t) \to \max \tag{4.9}$$

$$x_{t+1} = F_t(x_t, u_t) \quad \text{for all } t \in \mathbb{N}_0 \tag{4.10}$$

$$x_0 = \bar{x}. \tag{4.11}$$

Denote

$$\begin{aligned}
D_{x_t}F(\hat{x}_t, \hat{u}_t) &= A_t \quad \text{for all } t \in \mathbb{N}_0 \\
D_{u_t}F(\hat{x}_t, \hat{u}_t) &= B_t \quad \text{for all } t \in \mathbb{N}_0 \\
(A_0, A_1, ...) &= \mathbf{A} \\
(B_0, B_1, ...) &= \mathbf{B} \\
\mathcal{L} &= (\pi_0, (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})), \quad \mathcal{L} : \ell_\infty^n \times \ell_\infty^m \to \ell_\infty^n.
\end{aligned}$$

Again, we construct pertubations along the optimal solution, i.e. curves that start from the optimal solution, their directions are given and conditions (4.10), (4.11) are fulfilled.

Definition 4.1. We call a pair $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \ell_\infty^n \times \ell_\infty^m$ an admissible vector if there exist $\varepsilon_0 > 0$ and differentiable curves $\mathbf{p}(\varepsilon_0) = \{p_t(\varepsilon_0)\}_{t\in\mathbb{N}_0}$, $\mathbf{q}(\varepsilon_0) = \{p_t(\varepsilon_0)\}_{t\in\mathbb{N}_0}$, where

$$\begin{aligned}
p_t &: \langle 0, \varepsilon_0) \to X \\
q_t &: \langle 0, \varepsilon_0) \to U
\end{aligned}$$

for all $t \in \mathbb{N}_0$ such that the following conditions hold

i) $\mathbf{p}(0) = \mathbf{q}(0) = 0$

ii) $\mathbf{p}'(0) = \boldsymbol{\alpha}$ and $\mathbf{q}'(0) = \boldsymbol{\beta}$

iii) for each $\varepsilon \in \langle 0, \varepsilon_0)$ and $t \in \mathbb{N}_0$

$$\begin{aligned}
p_0(\varepsilon) &= 0 \\
\hat{x}_{t+1} + p_{t+1}(\varepsilon) &= F_t(\hat{x}_t + p_t(\varepsilon), \hat{u}_t + q_t(\varepsilon))
\end{aligned}$$

and $(\hat{\mathbf{x}} + \mathbf{p}(\varepsilon), \hat{\mathbf{u}} + \mathbf{q}(\varepsilon)) \in \ell_\infty^n \times \ell_\infty^m$.

If in any direction there exist an admissible perturbation curve, we can use it to derive the necessary conditions of optimality as in the case of linear autonomous dynamics. In the following proposition we state under which conditions this is the case. In order to prove it we apply implicit function theorem, bounded inverse theorem and following lemma.

Lemma 4.1. *Let X, Y be vector spaces, $T : X \to Y$ be a linear map and C be a closed complement of $\mathcal{N}(T)$ in X. Then the map $T : C \to \mathcal{R}(T)$ is an isomorphism. Futhermore, this map is also called a restriction of a map T to C and denoted by $T|_C$.*

Proof. We have to show that the map $T|_C$ is injective and surjective, so it is an isomorphism. Injectivity follows from

$$\mathcal{N}(T|_C) = \mathcal{N}(T) \cap C = \{0\}.$$

And as $X = C \oplus \mathcal{N}(T)$, one has $\mathcal{R}(T|_C) = T(C) = T(C \oplus \mathcal{N}(T)) = T(X) = \mathcal{R}(T)$. Hence it is surjective. □

Theorem 4.2. *(Bounded inverse theorem) Let X, Y be vector spaces, $T : X \to Y$ be a bounded linear operator that is one-to-one. Then the inverse map $T^{-1} : Y \to X$ is continuous.*

Proof. The theorem is in fact a corrolary to open mapping theorem and its proof can be found in Rudin [25]. □

Theorem 4.3. *(Implicit function theorem) Let X, Y, Z be Banach spaces, $U \subset X$, $V \subset Y$ open, $F : U \times V \to Z$ be C^r, $r \in (0, \infty\rangle$, $(x_0, y_0) \in U \times V$, $\Theta(x_0, y_0) = 0$. Let us assume that $D_{y_t}\Theta(x_0, y_0)$ has a continuous inverse operator. Then there exists a neighbourhoood $U_0 \times V_0 \subset U \times V$ of (x_0, u_0) and a function $\theta \in C^r(U_0, V_0)$ such that $\theta(x_0) = y_0$ and $\Theta(x, y) = 0$ for $(x, y) \in U_0 \times V_0$ hold if only if $y = \theta(x)$. Futhermore one has*

$$D\theta(x_0) = -[D_{y_t}\Theta(x_0, y_0)]^{-1} D_{x_t}\Theta(x_0, y_0).$$

Proof. The proof can be found in [24]. □

Proposition 4.2. Let us assume that $\mathcal{L}$ has a closed complement to its null space. Then each vector $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{N}(\mathcal{L})$ is admissible.

Proof. We apply implicit function theorem with $X = \mathbb{R}$, $Z = \ell_\infty^n$ and Y a closed complement to the null space of $\mathcal{L}$. Any pair $(\mathbf{x}, \mathbf{u}) \in \ell_\infty^{n+m}$ can be seperated to the sum of $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{N}(\mathcal{L})$ and $(\mathbf{a}, \mathbf{b}) \in Y$. We fix $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and construct curves $\mathbf{p}, \mathbf{q}$ such that

$$\begin{aligned} \mathbf{p}(\varepsilon) &= \varepsilon\boldsymbol{\alpha} + \mathbf{a}(\varepsilon), \quad \mathbf{a} : \langle 0, \varepsilon_0) \to Y, \\ \mathbf{q}(\varepsilon) &= \varepsilon\boldsymbol{\beta} + \mathbf{b}(\varepsilon), \quad \mathbf{b} : \langle 0, \varepsilon_0) \to Y \end{aligned}$$

and we prove that these curves fulfill the conditions (i)-(iii) from Definition 4.1. We define function $\Theta = (T_0, T) : X \times Y \to Z$ such that

$$\begin{aligned} T_0(\varepsilon, (\mathbf{a}, \mathbf{b})) &= (\varepsilon\alpha_0 + v_0 + \hat{x}_0) - \bar{x} \\ T(\varepsilon, (\mathbf{a}, \mathbf{b})) &= \boldsymbol{\sigma}(\varepsilon\boldsymbol{\alpha} + \mathbf{a}(\varepsilon) + \hat{\mathbf{x}}) - \mathbf{F}(\varepsilon\boldsymbol{\alpha} + \mathbf{a}(\varepsilon) + \hat{\mathbf{x}}, \varepsilon\boldsymbol{\beta} + \mathbf{b}(\varepsilon) + \hat{\mathbf{u}}), \end{aligned}$$

where $(\mathbf{F}(\mathbf{x}, \mathbf{u}))_t = F_t(x_t, u_t)$ and $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ is an optimal solution. Next, we set $(x_0, y_0) = (0, (\mathbf{0}_n, \mathbf{0}_m))$, where $\mathbf{0}_n = (0_n, 0_n, \dots)$, $\mathbf{0}_m = (0_m, 0_m, \dots)$ and prove that assumptions of the implicit function theorem are fulfilled.

As $(\hat{\mathbf{x}}, \hat{\mathbf{u}})$ fulfill state and initial conditions (4.10) - (4.11), we have

$$\begin{aligned} T_0(0, (\mathbf{0}_n, \mathbf{0}_m)) &= \hat{x}_0 - \bar{x} = 0_n \\ T(0, (\mathbf{0}_n, \mathbf{0}_m)) &= \boldsymbol{\sigma}\hat{\mathbf{x}} - \mathbf{F}(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \mathbf{0}_n. \end{aligned}$$

Since $F_t \in C^1$ it follows that the function $\Theta \in C^1$. In order to prove that $D_{(\mathbf{a},\mathbf{b})}\Theta(0, (\mathbf{0}_n, \mathbf{0}_m))$

has continuous inverse, we compute

$$D_{(\mathbf{a},\mathbf{b})}T_0(0,(\mathbf{0}_n,\mathbf{0}_m)) = \pi_0|_Y$$
$$D_{(\mathbf{a},\mathbf{b})}T(0,(\mathbf{0}_n,\mathbf{0}_m)) = (\boldsymbol{\sigma} - D_{\mathbf{x}}\mathbf{F}(\hat{\mathbf{x}},\hat{\mathbf{u}}) - D_{\mathbf{u}}\mathbf{F}(\hat{\mathbf{x}},\hat{\mathbf{u}}))|_Y = (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})|_Y,$$

So

$$D_{(\mathbf{a},\mathbf{b})}\Theta(0,(\mathbf{0}_n,\mathbf{0}_m)) = (\pi_0, (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B}))|_Y = \mathcal{L}|_Y$$

and by Lemma 4.1 it is an isomorphism as Y is a closed complement and by Theorem 4.2 it has a continuous inverse operator.

Hence according to the Implicit function theorem (Theorem 4.3) there exists a neighbourhood $X_0 \times Y_0 \subset X \times Y$ of $(0,(\mathbf{0}_n,\mathbf{0}_m))$ and a differentiable function $\theta : X_0 \to Y_0$, such that $\theta(\varepsilon) = (\mathbf{a}(\varepsilon), \mathbf{b}(\varepsilon))$ if only if

$$\theta(0) = (0,0)$$
$$T_0(\varepsilon,(\mathbf{a}(\varepsilon),\mathbf{b}(\varepsilon))) = 0$$
$$T(\varepsilon,(\mathbf{a}(\varepsilon),\mathbf{b}(\varepsilon))) = \mathbf{0}_n, \text{ for all } ((\varepsilon,(\mathbf{a}(\varepsilon),\mathbf{b}(\varepsilon)))) \in X_0 \times Y_0.$$

Since for all $\varepsilon \in X_0$ $(\mathbf{a}(\varepsilon),\mathbf{b}(\varepsilon)) \in Y_0 \subset \ell_\infty^n \times \ell_\infty^m$

$$\mathbf{p}(\varepsilon) + \hat{\mathbf{x}} = \mathbf{a}(\varepsilon) + \varepsilon\boldsymbol{\alpha} + \hat{\mathbf{x}} \in \ell_\infty^n$$
$$\mathbf{q}(\varepsilon) + \hat{\mathbf{x}} = \mathbf{b}(\varepsilon) + \varepsilon\boldsymbol{\beta} + \hat{\mathbf{u}} \in \ell_\infty^m,$$

So far we have proven that properties (i) and (iii) of an admissible vector are fulfilled, so it is left to prove (ii) that the direction of the curves are $(\boldsymbol{\alpha},\boldsymbol{\beta})$. Again, we apply implicit function theorem and compute

$$\theta'(0) = (\mathbf{a}'(0),\mathbf{b}'(0)) = -[D_{(\mathbf{a},\mathbf{b})}\Theta(0,(\mathbf{0}_n,\mathbf{0}_m))]^{-1}[D_\varepsilon\Theta(0,(\mathbf{0}_n,\mathbf{0}_m))].$$

So we have to find directional derivative $D_\varepsilon\Theta(0,(\mathbf{0}_n,\mathbf{0}_m))$.

$$\begin{aligned}
D_\varepsilon T_0(0,(\mathbf{0}_n,\mathbf{0}_m)) &= \alpha_0 = 0\\
D_\varepsilon T(0,(\mathbf{0}_n,\mathbf{0}_m)) &= (\boldsymbol{\sigma}\boldsymbol{\alpha} + v'(\varepsilon)\\
&\quad - D_{x_t}\mathbf{F}(\varepsilon\boldsymbol{\alpha} + \mathbf{a}(\varepsilon) + \hat{\mathbf{x}}, \varepsilon\boldsymbol{\beta} + \mathbf{b}(\varepsilon) + \hat{\mathbf{u}})(\boldsymbol{\alpha} + \mathbf{a}'(\varepsilon))\\
&\quad - D_{u_t}\mathbf{F}(\varepsilon\boldsymbol{\alpha} + \mathbf{a}(\varepsilon) + \hat{\mathbf{x}}, \varepsilon\boldsymbol{\beta} + \mathbf{b}(\varepsilon) + \hat{\mathbf{u}})(\boldsymbol{\beta} + \mathbf{b}'(\varepsilon)))|_{(\varepsilon,(\mathbf{a},\mathbf{b}))=(0,(\mathbf{0}_n,\mathbf{0}_m))}\\
&= \boldsymbol{\sigma}\boldsymbol{\alpha} - \mathbf{A}\boldsymbol{\alpha} - \mathbf{B}\boldsymbol{\beta} = (\boldsymbol{\sigma} - \mathbf{A}, -\mathbf{B})(\boldsymbol{\alpha},\boldsymbol{\beta})^\intercal
\end{aligned}$$

As $(\boldsymbol{\alpha},\boldsymbol{\beta})$ is in the null space of $\mathcal{L}$

$$D_\varepsilon\Theta(0,(\mathbf{0}_n,\mathbf{0}_m)) = \mathcal{L}(\boldsymbol{\alpha},\boldsymbol{\beta})^\intercal = (0,\mathbf{0}_n).$$

Therefore

$$\theta'(0) = (\mathbf{a}'(0),\mathbf{b}'(0)) = (\mathbf{0}_n,\mathbf{0}_m).$$

Finally we prove property (ii)

$$\begin{aligned}
\mathbf{p}'(0) &= \boldsymbol{\alpha} + \mathbf{a}'(0) = \boldsymbol{\alpha} + \mathbf{0}_n = \boldsymbol{\alpha}\\
\mathbf{q}'(0) &= \boldsymbol{\beta} + \mathbf{b}'(0) = \boldsymbol{\beta} + \mathbf{0}_m = \boldsymbol{\beta}.
\end{aligned}$$

□

Next, we proceed as in the problem with linear autonomous dynamics. If $(\boldsymbol{\alpha},\boldsymbol{\beta})$ is admissible and $(\hat{\mathbf{x}},\hat{\mathbf{u}})$ is an optimal solution then

$$\begin{aligned}
&\frac{\partial}{\partial\varepsilon}J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} \le 0\\
\frac{\partial}{\partial\varepsilon}J(\hat{\mathbf{x}} - \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} - \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} = -&\frac{\partial}{\partial\varepsilon}J(\hat{\mathbf{x}} + \varepsilon\boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon\boldsymbol{\beta})|_{\varepsilon=0} \le 0
\end{aligned} \tag{4.12}$$

Hence, again we have

$$
\begin{aligned}
0 &= \frac{\partial}{\partial \varepsilon} J(\hat{\mathbf{x}} + \varepsilon \boldsymbol{\alpha}, \hat{\mathbf{u}} + \varepsilon \boldsymbol{\beta})|_{\varepsilon=0} = \sum_{t=0}^{\infty} \delta^t [D_{x_t} f^0(\hat{x}_t, \hat{u}_t)\alpha_t + D_{u_t} f^0(\hat{x}_t, \hat{u}_t)\beta_t] \\
&= DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal .
\end{aligned}
$$

and the necessary conditions are analogous.

Theorem 4.4. *(Necessary conditions of optimality) Assume that the operator $\mathcal{L}$ has closed range. Then $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0$ for all admissible $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ if only if there exists $\boldsymbol{\phi} \in (\ell_\infty)^*$ such that*

$$DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}}) = \mathcal{L}^* \boldsymbol{\phi} \tag{4.13}$$

Moreover, if one has $\boldsymbol{\phi} = \boldsymbol{\psi} + \boldsymbol{\phi}^s$, where $\boldsymbol{\psi} = \{\psi_t\}_{t \in \mathbb{N}_0} \in \ell_1$ and $\boldsymbol{\phi}^s \in \ell_s$, then $DJ(\hat{\mathbf{x}}, \hat{\mathbf{u}})(\boldsymbol{\alpha}, \boldsymbol{\beta})^\intercal = 0$ for all admissible $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ if

$$
\begin{aligned}
D_{x_t} f^0(\hat{x}_t, \hat{u}_t) &= \psi_{t-1} - \delta A_t^* \psi_t \text{ for all } t \in \mathbb{N} \\
D_{u_t} f^0(\hat{x}_t, \hat{u}_t) &= -\delta B_t^* \psi_t \text{ for all } t \in \mathbb{N}_0 .
\end{aligned} \tag{4.14}
$$

Proof. As $\mathbf{A}$, $\mathbf{B}$ are general linear operators, by the Proposition 4.1 $\mathcal{L}$ is bounded. Next, we proceed analogously to the proof of the Theorem 4.1. □

Chapter 5

Closed Range of $\mathcal{L}$

In the previous chapter, we assumed that the operator $\mathcal{L} : \ell_\infty^n \times \ell_\infty^m \to \ell_\infty^n$, $\mathcal{L} = (\pi_0, (\sigma - \mathbf{A}, -\mathbf{B}))$ has closed range and that the complement to its null space exists and is closed as well. Now we show under which conditions this is the case. At first we explore the autonomous system, i.e. where matrices A_t, B_t are constant for any t, then we derive conditions for the nonautonomous system. In both cases we consider $\mathcal{L} : \ell_p^n \times \ell_p^m \to \ell_p^n$, $p \in \langle 1, \infty \rangle$.

5.1 Autonomous system

In this section we only consider matrix A that has no eigenvalues on the unit circle and prove that range of $\mathcal{L}$ and complement to null space of $\mathcal{L}$ are closed.

Proposition 5.1. If the eigenvalues of A do not lie on the unit circle, there exists a projection matrix P, such that $PA|_{\mathcal{R}(P)}$ has eigenvalues outside the unit circle, $(I - P)A|_{\mathcal{R}(I-P)}$ has eigenvalues inside the unit circle, hence it is regular. Moreover, there exist $C \geq 1$ and $\lambda \in (0, 1)$ such that

1. $\|(PA|_{\mathcal{R}(P)})^{t-s}\| < C\lambda^{t-s}, \quad$ for all $t \geq s$,

2. $\|((I - P)A|_{\mathcal{R}(I-P)})^{-(s-t)}\| < C\lambda^{s-t}, \quad$ for all $t < s$.

Proof. For the sake of simplicity, we drop the subscript and write PA instead of $PA|_{\mathcal{R}(P)}$ and QA insead of $(I-P)A|_{\mathcal{R}(I-P)}$. We denote by P the projection to the generalized eigenspaces of A corresponding to the eigenvalues inside the unit cirle, so $I-P=Q$ is the projection to the generalized eigenspaces of A corresponding to the eigenvalues outside the unit cirle. Then spectral radius $\rho(PA)$

$$\rho(PA) = \max\{|\lambda|, \lambda \in \mathrm{sp}(PA)\} < 1,$$

and

$$\min\{|\lambda|, \lambda \in \mathrm{sp}(QA)\} > 1.$$

So QA is a regular matrix and there exists its inverse and its spectral radius

$$\rho((QA)^{-1}) = \max\{|\lambda|, \lambda \in \mathrm{sp}((QA)^{-1})\} = \max\left\{\frac{1}{|\lambda|}, \lambda \in \mathrm{sp}(QA)\right\} < 1.$$

It is left to show that if $\rho(D) < 1$, then $\|D^t\| < c\mu^t$ for any $t \in \mathbb{N}$, where $c > 0$ and $\mu \in (0,1)$.

Any matrix D is similar to a matrix in Jordan canonical form J which has eigenvalues of D on its diagonal, 1 or 0 on the superdiagonal and zero everywhere else, i.e. there exists a matrix M_1 such that $D = M_1 J M_1^{-1}$.

Let us denote

$$J(\lambda,\varepsilon) = \begin{pmatrix} \lambda & \varepsilon & 0 & 0 & \dots & 0 \\ 0 & \lambda & \varepsilon & 0 & \dots & 0 \\ 0 & 0 & \lambda & \varepsilon & \dots & 0 \\ 0 & 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix},$$

so that $J(\lambda, 1)$ is a standard Jordan block for an eingenvalue λ. If λ is not a multiple eigenvalue, then $J(\lambda,\varepsilon) = (\lambda)$ for any $\varepsilon > 0$.

$J(\lambda, 1)$ is then similar to the matrix $J(\lambda, \varepsilon)$, where ε can be chosen arbitrary, i.e.

$$M(\lambda, \varepsilon)J(\lambda, 1) = J(\lambda, \varepsilon)M(\lambda, \varepsilon),$$

where

$$M(\lambda, \varepsilon) = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 0 & \frac{1}{\varepsilon} & \frac{1}{\varepsilon} & \frac{1}{\varepsilon} & \dots & \frac{1}{\varepsilon} \\ 0 & 0 & \frac{1}{\varepsilon^2} & \frac{1}{\varepsilon^2} & \dots & \frac{1}{\varepsilon^2} \\ 0 & 0 & 0 & \frac{1}{\varepsilon^3} & \dots & \frac{1}{\varepsilon^3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \frac{1}{\varepsilon^{n-1}} \end{pmatrix}$$

if λ is a multiple eigenvalue and $M = 1$ otherwise.

So there exists an invertible matrix M_ε such that

$$D = M_1 M_\varepsilon^{-1} J_\varepsilon M_\varepsilon M_1^{-1},$$

where J_ε is a block diagonal matrix with blocks $J(\lambda_i, \varepsilon)$ for each $\lambda_i \in \operatorname{sp}(D)$ and M_ε consists of the corresponding matrices $M(\lambda, \varepsilon)$.

Then

$$\begin{aligned} \|D^t\| = \|M_1 M_\varepsilon^{-1} J_\varepsilon^t M_\varepsilon M_1^{-1}\| &\leq \|M_1\| \|M_1^{-1}\| \|M_\varepsilon\| \|M_\varepsilon^{-1}\| \|J_\varepsilon^t\| \\ &\leq \|M_1\| \|M_1^{-1}\| \|M_\varepsilon\| \|M_\varepsilon^{-1}\| \|J_\varepsilon^t\|_1 \leq c \|J_\varepsilon^t\|_1. \end{aligned}$$

It is left to show that $\|J_\varepsilon^t\|_1$ is bounded by some μ^t, $\mu \in (0, 1)$.

By straightforward calculation

$$J(\lambda,\varepsilon)^t = \begin{pmatrix} \lambda^t & \binom{t}{1}\lambda^{t-1}\varepsilon & \binom{t}{2}\lambda^{n-1}\varepsilon^2 & \dots & \binom{t}{n-1}\lambda^{t-n+1}\varepsilon^{n-1} \\ 0 & \lambda^t & \binom{t}{1}\lambda^{t-1}\varepsilon & \dots & \binom{t}{m-2}\lambda^{t-m+2}\varepsilon^{n-2} \\ 0 & 0 & \lambda^t & \dots & \binom{t}{n-3}\lambda^{t-n+3}\varepsilon^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^t \end{pmatrix},$$

where for our convinience we define $\binom{t}{s} = 0$, if $s > t$. Hence if n is maximum multiplicity of the eigenvalues of D, then

$$\|J_\varepsilon^t\| \le \|J_\varepsilon^t\|_1 \le \sum_{i=0}^{\min\{t,n-1\}} \binom{t}{i} \varepsilon^i \left(\rho(D)\right)^{t-i} \le \sum_{i=0}^{t} \binom{t}{i} \varepsilon^i \left(\rho(D)\right)^{t-i} = (\varepsilon + \rho(D))^t.$$

If we choose ε sufficiently small, then $\mu = \varepsilon + \rho(D) < 1$ and

$$\|D^t\| \le c\|J_\varepsilon\|_1 \le c\mu^t.$$

As both PA and $(QA)^{-1}$ have eigenvalues inside the unit circle, they satisfy the following inequalities for any $t > s$

$$\|PA^{t-s}\| \le c_1\mu_1^{t-s}$$
$$\|(QA)^{s-t}\| \le c_2\mu_2^{t-s}.$$

Then we choose $C = \max\{1, c_1, c_2\}$, $\lambda = \max\{\mu_1, \mu_2\}$ so that we proved the claim. □

Below, we use the notation $A^- = PA|_{\mathcal{R}(P)}$ and $A^+ = (I-P)A|_{\mathcal{R}(I-P)}$, $B^- = PB$ and $B^+ = (I-P)B$. If $\mathbf{z} \in \ell_p^n$, then we may write $\mathbf{z} = \mathbf{z}^- \oplus \mathbf{z}^+$, where $\mathbf{z}^- = P\mathbf{z}$ and $\mathbf{z}^+ = (I-P)\mathbf{z}$. The dimension of $\mathcal{R}(P)$ is n_1 and the dimension of $\mathcal{R}(I-P)$ is n_2, such that $n = n_1 + n_2$. And we define $(A^-)^0 = I$. Next, we define

$$\Psi(t,s) = \begin{cases} (A^-)^{t-s} & \text{if } t \geq s \\ ((A^+)^{-1})^{s-t} & \text{if } t < s. \end{cases}$$

Note that for any $T \in \langle t, s\rangle$

$$\Psi(t,T)\Psi(T,s) = \Psi(t,s).$$

Theorem 5.1. *$\mathcal{L}$ has closed range and closed complement to its null space.*

Proof. In order to prove that $\mathcal{L}$ has closed range, we have to prove that the set of those $\mathbf{z} = \{z_t\}_{t\in\mathbb{N}_0}$ for which there exist $\mathbf{v} = \{v_t\}_{t\in\mathbb{N}_0}$, $\mathbf{w} = \{w_t\}_{t\in\mathbb{N}_0}$ such that

$$\begin{aligned} v_0 &= z_0 \\ v_{t+1} - Av_t - Bw_t &= z_t, \qquad \text{for all } t \in \mathbb{N} \end{aligned} \tag{5.1}$$

is closed. Applying our notation the equations (5.1) can be split into two sets of equations

$$\begin{aligned} v_0^- &= z_0^- & v_0^+ &= z_0^+ \\ v_{t+1}^- - A^-v_t^- - B^-w_t &= z_t^- & v_{t+1}^+ - A^+v_t^+ - B^+w_t &= z_t^+. \end{aligned}$$

By $\mathcal{L}^- : \ell_\infty^{n_1} \times \ell_\infty^m \to \ell_\infty^{n_1}$, $\mathcal{L}^+ : \ell_\infty^{n_2} \times \ell_\infty^m \to \ell_\infty^{n_2}$ denote the operators $\mathcal{L}^- = (\pi_0^-, \boldsymbol{\sigma}^- - \mathbf{A}^-, -\mathbf{B}^-)$ and $\mathcal{L}^+ = (\pi_0^+, \boldsymbol{\sigma}^+ - \mathbf{A}^+, -\mathbf{B}^+)$, where π_0^-, π_0^+ are operators π_0 defined on respective spaces and $\boldsymbol{\sigma}^-$, $\boldsymbol{\sigma}^+$ are $\boldsymbol{\sigma}$ defined on respective spaces. Note that $\mathcal{R}(\mathcal{L}) = \mathcal{R}(\mathcal{L}^-) \oplus \mathcal{R}(\mathcal{L}^+)$.

The first system leads to

$$
\begin{aligned}
v_0^- &= z_0^- \\
v_1^- &= A^- v_0^- + B^- w_0 + z_0^- = B^- w_0^- + A^- z_0^- + z_0^- \\
v_2^- &= A^- v_1^- + B^- w_1 + z_1^- = A^-(A^- z_0^- + B^- w_0 + z_0^-) + B^- w_1^- + z_1^- \\
&\vdots \\
v_t^- &= (A^-)^t z_0^- + \sum_{s=0}^{t-1} (A^-)^{t-s-1}(B^- w_s + z_s^-) = \Psi(t,0) z_0^- + \sum_{s=0}^{t-1} \Psi(t,s+1)(B^- w_s + z_s^-).
\end{aligned}
$$

And analogously it can be shown that for $t < T$ then

$$
\begin{aligned}
v_T^+ &= \Psi(T,t) v_t^+ + \sum_{s=t}^{T-1} \Psi(T,s+1)(B^+ w_s + z_s^+) \\
\Psi(T,t) v_t^+ &= v_T^+ - \sum_{s=t}^{T-1} \Psi(T,s+1)(B^+ w_s + z_s^+) \\
v_t^+ &= \Psi(t,T) v_T^+ - \sum_{s=t}^{T-1} \Psi(t,T)\Psi(T,s+1)(B^+ w_s + z_s^+) \\
v_t^+ &= \Psi(t,T) v_T^+ - \sum_{s=t}^{T-1} \Psi(t,s+1)(B^+ w_s + z_s^+)
\end{aligned}
$$

And for $T \to \infty$

$$
v_t^+ = -\sum_{s=t}^{\infty} \Psi(t,s+1)(B^+ w_s + z_s^+) \quad t \geq 1.
$$

In summary, if $\mathbf{z} = \mathcal{L}(\mathbf{v}, \mathbf{w})^\intercal$, then $\mathbf{v}, \mathbf{w} \in \ell_p^{n+m}$, $\mathbf{z} \in \ell_p^n$ and

$$v_t^- = \Psi(t,0)z_0^- + \sum_{s=0}^{t-1} \Psi(t,s+1)(B^- w_s + z_s^-) \quad t \geq 1 \tag{5.2}$$

$$v_0^- = z_0^- \tag{5.3}$$

$$v_t^+ = -\sum_{s=t}^{\infty} \Psi(t,s+1)(B^+ w_s + z_s^+) \quad t \geq 1 \tag{5.4}$$

$$v_0^+ = z_0^+ = -\sum_{s=0}^{\infty} \Psi(0,s+1)(B^+ w_s + z_s^+). \tag{5.5}$$

From the construction of the solutions $\mathbf{v}$ it is clear, that they are unique. Next we prove that for a given $\mathbf{z} \in \mathcal{R}(\mathcal{L})$ we obtain a unique solution $(\mathbf{v}, \mathbf{w})$ so that $\mathcal{R}(\mathcal{L})$ and $\mathcal{N}^C(\mathcal{L})$ are isomorphic. Then we show that they are closed.

Firstly, we show that for a given $\mathbf{z}^+ \in \ell_p^{n_2}$ there exists a solution $(\mathbf{v}^+, \mathbf{w}^+)$ such that (5.4) and (5.5) hold.

The space of all z_0^+ is n_2-dimensional, so

$$W = \left\{ \sum_{s=0}^{\infty} \Psi(0,s+1)B^+ w_s : \mathbf{w} \in \ell_p^m \right\}$$

has finite dimension $d < n_2$ as it is its subspace. Let us denote by $\xi_1, \ldots, \xi_d$, its basis vectors. Then there exist $\mathbf{w}^{(j)} = \{w_t^{(j)}\}_{t=0}^{\infty} \in \ell_p^n$ for all $j \in \{1, \ldots, d\}$ such that

$$\sum_{s=0}^{\infty} \Psi(0,s+1)B^+ w_s^{(j)} = \xi_j.$$

Then (5.5) is fulfilled if only if $z_0^+ + \sum_{s=0}^{\infty} \Psi(0,s+1)z_s^+ \in W$. Therefore for a given $\mathbf{z}^+ \in \mathcal{R}(\mathcal{L}^+)$ there exists $\alpha_1, \ldots, \alpha_d$ such that $z_0^+ + \sum_{s=0}^{\infty} \Psi(0,s+1)z_s^+ = \sum_{j=1}^{d} \alpha_j \xi_j$. Denote the space of such $\mathbf{z}^+ \in Z^+ \subset \ell_p^{n_2}$. Then

$$\begin{aligned}
z_0^+ + \sum_{s=0}^{\infty} \Psi(0,s+1) z_s^+ &= \sum_{j=1}^{d} \alpha_j \xi_j \\
&= \sum_{j=1}^{d} \alpha_j \sum_{s=0}^{\infty} \Psi(0,s+1) B^+ w_s^{(j)} \\
&= \sum_{s=0}^{\infty} \Psi(0,s+1) \sum_{j=1}^{d} \alpha_j w_s^{(j)}
\end{aligned}$$

And so we set

$$\begin{aligned}
\mathbf{w} &= \sum_{j=1}^{d} \alpha_j \mathbf{w}^{(j)} \\
v_0^+ &= z_0^+ \\
v_t^+ &= -\sum_{s=t}^{\infty} \Psi(t,s+1)(z_s^+ + B^+ w_s) \quad t \geq 1.
\end{aligned}$$

And for a given $\mathbf{z}^- \in \mathcal{R}(\mathcal{L}^-)$ we obtain

$$\begin{aligned}
v_0^- &= z_0^- \\
v_t^- &= \Psi(t,0) z_0^- + \sum_{s=0}^{t-1} \Psi(t,s+1)(z_s^- + B^- w_s), \quad t \geq 1.
\end{aligned}$$

As $\mathbf{w}$ is a linear combination of $\mathbf{w}^{(j)} \in \ell_p^m$ it belongs to ℓ_p^m as well. Next, we show that $\mathbf{v} = \mathbf{v}^- \oplus \mathbf{v}^+ \in \ell_p^n$. For $p \in \langle 1, \infty)$ we obtain

$$\begin{aligned}
|v_t^-|^p &= \left| (\Psi(t,0) + \Psi(t,1))(z_0^- + B^- w_0) + \sum_{s=1}^{t-1} \Psi(t,s+1) \left(z_s^- + B^- w_s \right) \right|^p \\
&\leq \left((C\lambda^t + C\lambda^{t-1}) |z_0^- + B^- w_0| + \sum_{s=1}^{t-1} C\lambda^{t-s-1} |z_s^- + B^- w_s| \right)^p \\
&\leq C^p \left(2\lambda^{t-1} |z_0^- + B^- w_0| + \sum_{s=1}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s| \right)^p \\
&\leq (2C)^p \left(\sum_{s=0}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s| \right)^p, \quad t \geq 1.
\end{aligned}$$

and

$$\begin{aligned} |v_t^+|^p &= \left| \sum_{s=t}^{\infty} \Psi(t, s+1)(z_s + B^+ w_s) \right|^p \\ &\leq C^p \left(\sum_{s=t}^{\infty} \lambda^{s-t+1} |z_s + B^+ w_s| \right)^p . \end{aligned}$$

In case $p > 1$ we apply Jensen's inequality

$$\left(\frac{\sum_{s=0}^{\infty} \lambda^{t-s-1} |z_s^-|}{\sum_{s=0}^{\infty} \lambda^{t-s-1}} \right)^p \leq \frac{\sum_{s=0}^{\infty} \lambda^{t-s-1} |z_s^-|^p}{\sum_{s=0}^{\infty} \lambda^{t-s-1}}.$$

Hence, we have

$$\begin{aligned} \|\mathbf{v}^-\|_p \leq \left(\sum_{t=1}^{\infty} |v_t^-|^p \right)^{\frac{1}{p}} &\leq 2C \left(\sum_{t=1}^{\infty} \left(\sum_{s=0}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s| \right)^p \right)^{\frac{1}{p}} \\ &\leq 2C \left(\sum_{t=1}^{\infty} \left(\sum_{s=0}^{t-1} \lambda^{t-s-1} \right)^{p-1} \sum_{s=0}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s|^p \right)^{\frac{1}{p}} \\ &= 2C \left(\sum_{t=1}^{\infty} \left(\frac{1-\lambda^t}{1-\lambda} \right)^{p-1} \sum_{s=0}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s|^p \right)^{\frac{1}{p}} \\ &\leq 2C \left(\frac{1}{1-\lambda} \right)^{\frac{p-1}{p}} \left(\sum_{t=1}^{\infty} \sum_{s=0}^{t-1} \lambda^{t-s-1} |z_s^- + B^- w_s|^p \right)^{\frac{1}{p}} \\ &\leq 2C \left(\frac{1}{1-\lambda} \right)^{\frac{p-1}{p}} \left(\sum_{s=0}^{\infty} |z_s^- + B^- w_s|^p \sum_{t=s+1}^{\infty} \lambda^{t-s-1} \right)^{\frac{1}{p}} \\ &\leq \frac{2C}{1-\lambda} \left(\sum_{s=0}^{\infty} |z_s^- + B^- w_s|^p \right)^{\frac{1}{p}} \end{aligned}$$

Next, we apply Minkowski's inequality

$$\begin{aligned}
\|\mathbf{v}\|_p &\leq \frac{2C}{1-\lambda}\left(\sum_{s=0}^{\infty}|z_s^- + B^- w_s|^p\right)^{\frac{1}{p}} \\
&\leq \frac{2C}{1-\lambda}\left(\|\mathbf{z}\|_p + \|B^-\|_p\|\mathbf{w}\|_p\right) < \infty
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{v}^+\|_p = \left(\sum_{t=0}^{\infty}|v_t^+|^p\right)^{\frac{1}{p}} &\leq C\left(\sum_{t=0}^{\infty}\left(\sum_{s=t}^{\infty}\lambda^{s-t+1}|z_s^+ + B^+ w_t|\right)^p\right)^{\frac{1}{p}} \\
&\leq C\left(\sum_{t=0}^{\infty}\left(\frac{\lambda}{1-\lambda}\right)^{p-1}\sum_{s=t}^{\infty}\lambda^{s-t+1}|z_s^+ + B^+ w_s|^p\right)^{\frac{1}{p}} \\
&= C\left(\frac{\lambda}{1-\lambda}\right)^{\frac{p-1}{p}}\left(\sum_{t=0}^{\infty}\sum_{s=t}^{\infty}\lambda^{s-t+1}|z_s^+ + B^+ w_s|^p\right)^{\frac{1}{p}} \\
&= C\left(\frac{\lambda}{1-\lambda}\right)^{\frac{p-1}{p}}\left(\sum_{s=0}^{\infty}|z_s^+ + B^+ w_s|^p\sum_{t=0}^{s}\lambda^{s-t+1}\right)^{\frac{1}{p}} \\
&\leq C\left(\frac{\lambda}{1-\lambda}\right)^{\frac{p-1}{p}}\left(\sum_{s=0}^{\infty}|z_s^+ + B^+ w_t|^p\sum_{t=1}^{\infty}\lambda^{t}\right)^{\frac{1}{p}} \\
&\leq \frac{C\lambda}{1-\lambda}\left(\left(\sum_{s=0}^{\infty}|z_s^+|^p\right)^{\frac{1}{p}} + \left(\sum_{s=0}^{\infty}|B^+ w_t|^p\right)^{\frac{1}{p}}\right) \\
&\leq \frac{C\lambda}{1-\lambda}\left(\|\mathbf{z}^+\|_p + \|B^+\|_p\|\mathbf{w}\|_p\right) < \infty.
\end{aligned}$$

In case $p = 1$ we obtain

$$\begin{aligned}
\|\mathbf{v}^-\|_1 &\leq 2C\sum_{t=1}^{\infty}\sum_{s=0}^{t-1}\lambda^{t-s-1}|z_s^- + B^- w_s| \leq 2C\sum_{s=0}^{\infty}|z_s^- + B^- w_s|\sum_{t=s-1}^{\infty}\lambda^{t-s-1} \\
&\leq \frac{2C}{1-\lambda}\left(\|\mathbf{z}^-\|_1 + \|B^-\|_1\|\mathbf{w}\|_1\right) < \infty
\end{aligned}$$

and

$$\begin{aligned}\|\mathbf{v}^+\|_1 &\leq C\sum_{t=0}^{\infty}\left(\sum_{s=t}^{\infty}\lambda^{s-t+1}|z_s^+ + B^+w_s|\right) \leq C\sum_{s=0}^{\infty}|z_s^+ + B^+w_t|\sum_{t=1}^{\infty}\lambda^t \\ &\leq \frac{C\lambda}{1-\lambda}\left(\|\mathbf{z}^+\|_1 + \|B^+\|_1\|\mathbf{w}\|_1\right) < \infty.\end{aligned}$$

Finally, for $p = \infty$ we have

$$\begin{aligned}\|\mathbf{v}^-\|_\infty &= \sup_{t\in\mathbb{N}}|v_t^-| \leq 2C\sup_{t\in\mathbb{N}}\sum_{s=0}^{t-1}\lambda^{t-s-1}|z_s^- + B^-w_s| \leq 2C\sup_{t\in\mathbb{N}}|z_t^- + B^-w_s|\sup_{t\in\mathbb{N}}\sum_{s=0}^{t-1}\lambda^{t-s-1} \\ &\leq \frac{2C}{1-\lambda}\left(\|\mathbf{z}^-\|_\infty + \|B^-\|_\infty\|\mathbf{w}\|_\infty\right) < \infty.\end{aligned}$$

$$\begin{aligned}\|\mathbf{v}^+\|_\infty &= \sup_{t\in\mathbb{N}}|v_t^+| \leq C\sup_{t\in\mathbb{N}}\sum_{s=t}^{\infty}\lambda^{s-t+1}|z_s^+ + B^+w_t| \leq C\sup_{t\in\mathbb{N}}|z_t^+ + B^+w_t|\sup_{t\in\mathbb{N}}\sum_{s=t}^{\infty}\lambda^{s-t+1} \\ &\leq \frac{C\lambda}{1-\lambda}\left(\|\mathbf{z}^+\|_\infty + \|B^+\|_\infty\|\mathbf{w}\|_\infty\right) < \infty.\end{aligned}$$

We proved that $\mathbf{v} = \mathbf{v}^- \oplus \mathbf{v}^+ \in \ell_p^n$, it is unique for a given $\mathbf{z}, \mathbf{w}$ and $\mathbf{w} \in \ell_p^m$ is unique for a given $\mathbf{z} \in Z^+$, hence $\mathcal{L}$ is one-to-one and its range is isomorphic to complement to its null space. Futhermore, we have shown that $\mathcal{R}(\mathcal{L}^+) = Z^+$ and $\mathcal{R}(\mathcal{L}^-) = \ell_p^{n_1}$.

The space Z^+ has finite codimension, so it is closed as well as $\ell_p^{n_1}$. Hence $\mathcal{R}(\mathcal{L})$ is closed and due to their isomorphism $\mathcal{N}^C(\mathcal{L})$ is closed as well.

This completes the proof.

□

5.2 Nonautonomous system

Definition 5.1. (Exponential Dichotomy)

Let us consider a linear difference equation

$$v_{t+1} = A_t v_t + B_t w_t + z_t, \tag{5.6}$$

with an initial condition $v_0 = z_0$ where $\mathbf{v} = \{v_t\}_{t\in\mathbb{N}} \in \ell_p^n$, $\mathbf{w} = \{w_t\}_{t\in\mathbb{N}} \in \ell_p^m$ $p \in \langle 1, \infty\rangle$,

$t \in \mathbb{N}$ and A_t are $n \times n$ matrices. We say that the linear difference equation (5.6) has an exponential dichotomy on $\mathbb{N}$ if there exist $C \geq 1$, $\lambda \in (0,1)$ and a family of projections $P_t, t \in \mathbb{N}$ such that

1. $P_{t+1}A_t = A_t P_t$, i.e. they commute,

2.
$$\|P_t\| \leq C \tag{5.7}$$
$$\left\| \prod_{i=t-1}^{s} A_j^- \right\| \leq C\lambda^{t-s} \text{ for all } t \geq s, \tag{5.8}$$

   where $A_t^- = P_{t+1}A_t|_{\mathcal{R}(P_t)}$,

3. $A_j^+ = (I - P_{t+1})A_t|_{\mathcal{R}(I-P_t)} = Q_{t+1}A_t|_{\mathcal{R}(Q_t)}$ are invertible for all $t \in \mathbb{N}$ and
$$\left\| \prod_{i=t}^{s-1} (A_j^+)^{-1} \right\| \leq C\lambda^{s-t}, \text{ for all } t < s. \tag{5.9}$$

For the sake of simplicity, let us denote

$$\Psi(t,s) = \begin{cases} \prod_{i=t-1}^{s} A_j^-, & \text{if } t \geq s \\ \prod_{i=t}^{s-1} (A_j^+)^{-1} & \text{if } t < s. \end{cases}$$

so that equations (5.8), (5.9) can be rewritten to

$$\|\Psi(t,s)\| \leq C\lambda^{|t-s|}.$$

Note that in case A_t are constant for all $t \in \mathbb{N}_0$,

$$\Psi(t,s) = \begin{cases} (A^-)^{t-s}, & \text{if } t \geq s \\ (A^+)^{t-s} & \text{if } t < s. \end{cases}$$

Theorem 5.2. *Let the linear difference equation (5.6) have an exponential dichotomy on* $\mathbb{N}$ *and* B_t *be bounded for all* t*. Then the operator* $\mathcal{L}$ *has closed range.*

Proof. Assume that linear difference equation (5.6) has an exponential dichotomy with constants C, λ and a family of projection matrices P_t, $I - P_t = Q_t, t \in \mathbb{N}$ and let $\mathbf{z} \in \mathcal{R}(\mathcal{L})$. As v_t can be rewritten as $(P_t + Q_t)v_t = P_t v_t + Q_t v_t$ the system (5.6) leads to two systems of equations

$$\begin{aligned} P_0 v_0 &= P_0 z_0 \\ P_{t+1} v_{t+1} &= P_{t+1} A_t v_t + P_{t+1} B_t w_t + P_{t+1} z_t \end{aligned} \tag{5.10}$$

$$\begin{aligned} Q_0 v_0 &= Q_0 z_0 \\ Q_{t+1} v_{t+1} &= Q_{t+1} A_t v_t + Q_{t+1} B_t w_t + Q_{t+1} z_t. \end{aligned} \tag{5.11}$$

As Q_t are projection matrices and they commute, i.e. $Q_{t+1}A_t = Q_{t+1}^2 A_t = Q_{t+1} A_t Q_t$, the last equation (5.11) can be rewritten to

$$\begin{aligned} Q_{t+1} v_{t+1} &= Q_{t+1} A_t Q_t v_t + Q_{t+1} B_t w_t + Q_{t+1} z_t \\ Q_t v_t &= (Q_{t+1} A_t)^{-1} (Q_{t+1} v_{t+1} - Q_{t+1} B_t w_t - Q_{t+1} z_t) \end{aligned}$$

Denote $P_{t+1} z_t = z_t^-$, $Q_{t+1} z_t = z_t^+$, $P_t v_t = v_t^-$, $Q_t v_t = v_t^+$ and $P_{t+1} B_t = B_t^-$, $Q_{t+1} B_t = B_t^+$ for any $t \in \mathbb{N}$, so that we may write

$$\begin{aligned} v_0^- &= z_0^- & v_0^+ &= z_0^+ \\ v_{t+1}^- - A_t^- v_t^- - B_t^- w_t^- &= z_t^- & v_t^+ &= (A_t^+)^{-1} (v_{t+1}^+ - B_t^+ w_t^+ - z_t^+). \end{aligned}$$

Then again as in the Proposition 5.1 it can be inductively shown that

$$v_t^- = \Psi(t,0)z_0^- + \sum_{s=0}^{t-1} \Psi(t,s+1)(z_s^- + B_s^- w_s)$$
$$v_t^+ = -\sum_{s=t}^{\infty} \Psi(t,s+1)(z_s^+ + B_s^+ w_s)$$
$$v_0^+ = z_0^+ = -\sum_{s=0}^{\infty} \Psi(0,s+1)(z_s^+ + B_s^+ w_s). v_0^- = z_0^-$$

For the rest of the proof we proceed as in the proof of the Proposition 5.1 applying inequalities

$$\sum_{s=0}^{\infty} |B_s w_s|^p \le \sup_{s\ge 0} \|B_s\|^p \sum_{s=0}^{\infty} |w_s|^p, \qquad \text{for } p \in \langle 0, \infty)$$

and for $p = \infty$

$$\sup_{s\ge 0} |B_s w_s| \le \sup_{s\ge 0} \|B_s\| \sup_{s\ge 0} |w_s| = \|\mathbf{w}\|_\infty \sup_{s\ge 0} \|B_s\|.$$

□

5.3 Special Cases

Proposition 5.2. If matrices A_t converge to a matrix A_∞ such that its eigenvalues do not lie on the unit circle, the linear difference equation (5.6) has an exponential dichotomy on $\mathbb{N}$.

Proof. In progress. □

Proposition 5.3. If matrices A_t are periodic with period T and the matrix $\mathcal{A} = A_T A_{T-1} \dots A_1$ has no eigenvalues on unit cirle and it is regular, then the linear difference equation (5.6) has exponential dichotomy on $\mathbb{N}$.

Proof. Let A_t be periodic with minimum period T, i.e. $A_{t+T} = A_t$ for all $t \in \mathbb{N}$ and let

$$\mathcal{A}_k = A_{(1+k)T} A_{(1+k)T-1} \dots A_{kT+1}, k \in \mathbb{N}_0,$$

then $\mathcal{A}_k$ is constant for any k.

We can rewrite the linear difference equations into

$$\begin{aligned} v_{t+1} &= A_t v_t + B_t w_t \\ v_{t+2} &= A_{t+1}A_t v_t + A_{t+1}B_t w_t + B_{t+1}w_{t+1} \\ &\vdots \\ v_{t+T+1} &= A_{t+T}\dots A_{t+1}A_t v_t + \sum_{s=t}^{t+T-1} A_{t+T}\dots A_{s+1}(B_s w_s) + B_{t+T}w_{t+T}. \end{aligned}$$

If we denote

$$\begin{aligned} \nu_k &= v_{kT+1} \\ \xi_k &= (B_{1+(k-1)T}w_{1+(k-1)T}, B_{2+(k-1)T}w_{2+(k-1)T}, \dots, B_{1+kT}w_{1+kT})^\top \\ \mathcal{B} &= (A_{1+kT}A_{kT}\dots A_{2+(k-1)T}, A_{2+kT}A_{kT}\dots A_{3+(k-1)T}, \dots, A_{1+kT}, I) \end{aligned}$$

for any $k \in \mathbb{N}$, then the system can be rewritten to an autonomous system

$$\nu_{k+1} = \mathcal{A}\nu_k + \mathcal{B}\xi_k.$$

According to Proposition 5.1, there exist projection matrices P, $Q = I - P$ and $\lambda \in (0,1), C_{\mathcal{A}} \geq 1$ such that

$$\|P\mathcal{A}|^{t-s}_{\mathcal{R}(P)}\| < C_{\mathcal{A}}\lambda^{t-s}, \text{ if } t \geq s \tag{5.12}$$

$$\|Q\mathcal{A}|^{t-s}_{\mathcal{R}(Q)}\| < C_{\mathcal{A}}\lambda^{s-t}, \text{ if } t < s. \tag{5.13}$$

Again, we will leave out the subscripts $\mathcal{R}(P)$ and $\mathcal{R}(Q)$. As $\mathcal{A}$ is regular, all the matrices A_t are regular and we can define a family of matrices

$$P_t = \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right), \text{ for all } t \geq 1.$$

Then for all $t \geq 1$

$$\begin{aligned} Q_t &= \left(\prod_{i=t-1}^{1} A_i\right) Q \left(\prod_{i=1}^{t-1} A_i^{-1}\right) \\ &= \left(\prod_{i=t-1}^{1} A_i\right) I \left(\prod_{i=1}^{t-1} A_i^{-1}\right) - \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right) = I - P_t \end{aligned}$$

and P_t are projection matrices

$$\begin{aligned} P_t^2 &= \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right) \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right) = \left(\prod_{i=t-1}^{1} A_i\right) P^2 \left(\prod_{i=1}^{t-1} A_i^{-1}\right) \\ &= \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right) = P_t, \end{aligned}$$

they commute

$$P_{t+1} A_t = \left(\prod_{i=t}^{1} A_i\right) P \left(\prod_{i=1}^{t} A_i^{-1}\right) A_t = A_t \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t} A_i^{-1}\right) A_t^{-1} A_t = A_t P_t$$

and they are periodic with period T

$$\begin{aligned} P_{t+T} &= \left(\prod_{i=t+T-1}^{1} A_i\right) P \left(\prod_{i=1}^{t+T-1} A_i^{-1}\right) = \left(\prod_{i=t+T-1}^{T+1} A_i\right) \mathcal{A} P \mathcal{A}^{-1} \left(\prod_{i=T+1}^{t+T-1} A_i^{-1}\right) \\ &= \left(\prod_{i=t-1}^{1} A_i\right) P \mathcal{A} \mathcal{A}^{-1} \left(\prod_{i=1}^{t-1} A_i^{-1}\right) = \left(\prod_{i=t-1}^{1} A_i\right) P \left(\prod_{i=1}^{t-1} A_i^{-1}\right) = P_t. \end{aligned}$$

Hence

$$\|P_t\| \leq \max_{i \in \{1,2,\ldots,T\}} \|P_i\| = C_P.$$

It is left to show that for any $t, s \in \mathbb{N}, t > s$

$$\left\| \prod_{i=t-1}^{s} A_i^- \right\| < C\lambda^{t-s} \tag{5.14}$$

$$\left\| \prod_{i=s}^{t-1} (A_i^+)^{-1} \right\| < C\lambda^{t-s}, \tag{5.15}$$

where $\lambda \in (0,1)$ and $C \geq 1$.

We split the proof into two parts.

1. Let us assume that $t - s > T$.

   As projection matrices P_t commute, $\|P_{t-1}A_{t-1} \dots A_s\| = \|A_{t-1} \dots P_{k+1}A_k \dots A_s\|$ for any $k \in \langle s+1, t) \cap \mathbb{N}$.

   Denote $\alpha = \lfloor \frac{t}{T} \rfloor - \lceil \frac{s}{T} \rceil \geq t - s - 2T$, then

$$\begin{aligned}
\left\| \prod_{i=t-1}^{s} A_i^- \right\| &= \|P_t A_{t-1} \dots A_s\| \\
&= \|A_{t-1} \dots A_{\lfloor \frac{t}{T} \rfloor T+1} P \mathcal{A}^\alpha A_{\lceil \frac{s}{T} \rceil T} \dots A_s\| \\
&\leq \|A_{t-1}\| \dots \|A_{\lfloor \frac{t}{T} \rfloor T+1}\| \|P\mathcal{A}^\alpha\| \|A_{\lceil \frac{s}{T} \rceil T}\| \dots \|A_s\| \leq C_\Pi^2 C_\mathcal{A} \lambda^\alpha \\
&\leq C_\Pi^2 C_\mathcal{A} \mu^{t-s-2T} = C_\Pi^2 C_\mathcal{A} \lambda^{-2T} \lambda^{t-s},
\end{aligned}$$

   where $C_\Pi = \|A_T\|\|A_{T-1}\| \dots \|A_1\|$.

2. Let $0 < t - s \leq T$, then

$$\begin{aligned}
\left\| \prod_{i=t-1}^{s} A_i^- \right\| &\leq \|P_t\|\|A_{t-1}\| \dots \|A_s\| \\
&\leq C_P C_\Pi = C_P \frac{C_\Pi}{\lambda^{t-s}} \lambda^{t-s} \leq C_P \frac{C_\Pi}{\lambda^T} \lambda^{t-s}
\end{aligned}$$

   If we put $C = \max\{1, C_\Pi^2 C_\mathcal{A} \lambda^{-2T}, C_\Pi C_P \lambda^{-T}\}$, we obtain the required inequality.

The proof for the inequality (5.15) is analogous. □

Proposition 5.4. If the linear difference equation (5.6) has an exponential dichotomy on $\mathbb{N}\backslash K$, where $K = \{1, \ldots, T\}$, then it has exponential dichotomy on $\mathbb{N}$.

Proof. Let (5.6) have an exponential dichotomy on $\mathbb{N}\backslash K$ with projections P_t, Q_t constants $C_1 \geq 1$ and $\lambda \in (0,1)$. Next, we define

$$\mathcal{R}(P_t) = \left\{ v \in \mathbb{R}^n : \prod_{i=T-1}^{s} A_t v \in \mathcal{R}(P_T) \right\}$$
$$\mathcal{R}(Q_t) = \left\{ v \in \mathbb{R}^n : \prod_{i=t}^{T-1} A_t^{-1} v \in \mathcal{R}(Q_T) \right\}.$$

Next, for any $t, s < T$ we define

$$\Psi(t,s) = \begin{cases} \prod_{i=t-1}^{s} A_t|_{\mathcal{R}(P_t)}, & \text{if } t \geq s \\ \prod_{i=t}^{s-1} A_t^{-1}|_{\mathcal{R}(Q_t)} & \text{if } t < s. \end{cases}$$

Let $C_2 = \sup_{t \in \langle 0, T \rangle} \{\|\Psi(t,T)\|, \|\Psi(T,t)\|\}$. Then

$$\begin{aligned} \|\Psi(t,s)\| &\leq \|\Psi(t,T)\| \|\Psi(T,s)\| \leq C_1 C_2 \lambda^{t-T} \\ &\leq C_1 C_2 \lambda^{s-T} \lambda^{t-s} \leq C_1 C_2 \lambda^{-T} \lambda^{t-s}, \text{ if } s \leq T \leq t \\ \|\Psi(t,s)\| &\leq \|\Psi(t,T)\| \|\Psi(T,s)\| \leq C_1 C_2 \lambda^{s-T} \\ &\leq C_1 C_2 \lambda^{t-T} \lambda^{s-t} \leq C_1 C_2 \lambda^{-T} \lambda^{s-t}, \text{ if } t \leq T \leq s. \end{aligned}$$

Hence the equation (5.6) has exponential dichotomy on $\mathbb{N}$ with $C = \max\{C_1, C_1 C_2 \lambda^{-T}\}$ and $\lambda \in (0,1)$.

□

Finally, in the following example we show that if the system (5.6) does not possess exponential dichotomy, $\mathcal{L}$ may not have closed range.

Example 1. We assume linear autonomous dynamics and $A = 1$, $B = 0$, so that the state can be rewritten as $x_{t+1} = x_t$. Then

$$\mathcal{R}(\mathcal{L}) = \{\mathbf{z} \in \ell^1_\infty : z_t = x_{t+1} - x_t, \mathbf{x} \in \ell^1_p\}.$$

For a given $\varepsilon > 0$, we choose $\mathbf{z}^\varepsilon$ such that $z^\varepsilon_0 = 0$ and $z^\varepsilon_t = t^{-(1+\varepsilon)}$ for $t \geq 1$. Then corresponding $\mathbf{x}^\varepsilon$ is given as $x^\varepsilon_0 = 0$ and $x^\varepsilon_t = \sum_{s=2}^{t-1}(s-1)^{-(1+\varepsilon)}$, $t \geq 1$. So while $\mathbf{z}^\varepsilon, \mathbf{x}^\varepsilon \in \ell^1_\infty$ and $\lim\limits_{\varepsilon\to 0} z_t{}^\varepsilon = t^{-1}$, so $\lim\limits_{\varepsilon\to 0} \mathbf{z}^\varepsilon \in \ell^1_\infty$, but $\lim\limits_{\varepsilon\to 0} \mathbf{x}^\varepsilon \notin \ell^1_\infty$. Therefore $\mathcal{R}(\mathcal{L})$ is not closed.

Conclusion

We focused on infinite-horizon, discrete-time optimal control problems and established necessary conditions of optimality of Potryagin maximum principle type. We considered problems with linear autonomous dynamics $x_{t+1} = Ax_t + Bu_t + d$ and general dynamics $x_{t+1} = F_t(x_t, u_t)$.

Current literature associated with our research gave us two main results. Blot and Chebbi [5] established the maximum principle in the space ℓ_1 with objective function without discount by reduction to finite horizon and imposed the condition that $A_t = D_{x_t}F_t(\hat{x}_t, \hat{u}_t)$ are invertible for all t. Later, Blot and Hayek [6] considered the same problem as we did and via tools of functional analysis formulated condition

$$\sup_{t \in \mathbb{N}_0} \|A_t\|_\infty < 1.$$

We also employed direct approach rather than reduction to finite horizon and faced four main challenges. At first, we had to prove that the objective function is Fréchet differentiable. Then by standard method of constructing perturbations along the optimal solution, we derived necessary conditions of optimality with adjoint variable belonging to the dual space of ℓ_∞, $(\ell_\infty)^* = \ell_1 \oplus \ell_s$. We managed to circumvent its non-sequential component ℓ_s. The most significant results are described in the last chapter where we formulate assumptions under which the necessary conditions hold, i.e. when the operator $\mathcal{L}$ has closed range. Moreover, in the case of general dynamics we had to show that its null space is complemented and the complement is closed.

In the case of linear autonomous dynamics $x_{t+1} = Ax_t + Bu_t + d$, it is sufficient that A has no eigenvalues on the unit circle. In case of general dynamics, the assumption is formulated as exponential dichotomy, i.e. there exist $C \geq 1$, $\lambda \in (0,1)$ and bounded projection matrices P_t such that

$$\|\Psi(t,s)\| \leq C\lambda^{|t-s|}, \quad \text{for any } t,s \in \mathbb{N}_0,$$

where

$$\Psi(t,s) = \begin{cases} \prod_{i=t-1}^{s} P_{t+1}A_t|_{\mathcal{R}(P_t)}, & \text{if } t \geq s \\ \prod_{i=t}^{s-1}(I-P_{t+1})A_t^{-1}|_{\mathcal{R}(I-P_t)} & \text{if } t < s. \end{cases}$$

In comparison to the previous results, we managed to formulate the maximum principle with $\psi_0 = 1$. In Blot, Hayek [6] they proved that their ψ_0 is non-zero, however Blot, Chebbi [5] did not exclude this possibility.

In finite horizon problems without terminal constraints the transversality condition is $\psi_T = 0$. In our case, we do not have a terminal condition as our response is from ℓ_∞ and we obtain the condition $\boldsymbol{\psi} \in \ell_1$, i.e. $\sum_{t=1}^{\infty} |\psi_t| < \infty$ which might be understood as a transversality condition in case of the infinite horizon problems.

While the condition of exponential dichotomy extends the framework of problems for which the maximum principle hold, we could not formulate it as an equivalence. However, we found an example where exponential dichotomy is not satisfied and the range of $\mathcal{L}$ is not closed, hence the closed range theorem cannot be applied.

The thesis ends with the examples of systems having exponential dichotomy, but there is definitely still a lot of space for future development of the presented framework. Moreover, further research can also be conducted in order to examine necessity of condition of exponential dichotomy.

Bibliography

[1] PONTRYAGIN, Lev et al. *The Mathematical Theory of Optimal Processes* . New York: Interscience Publishers, 1962.

[2] ARAUJO, A. & SCHEINKMAN & J. A. *Smoothness, comparative dynamics, and the turnpike property.* In: Econometrica, vol. 45, no. 3, 1977, pp. 601–620.

[3] MARCUS M. & ZASLAVSKI A.J. *The Structure of Extremals of a Class of Second Order Variational Problems.* Ann. Inst. H. Poincaré, Anal. non linéaire, 16: 593-629, 1999.

[4] ANDERSON, B.D.O. & MOORE, J.B. *Linear Optimal Control.* Prentice-Hall, Englewood Cliffs, NJ, 1971.

[5] BLOT, J. & CHEBBI, H. *Discrete Time Pontryagin Principles with Infinite Horizon.* In: Journal of Mathematical Analysis & Applications, vol. 246, 2000, pp. 265-279.

[6] BLOT, J. & HAYEK, N. *Infinite Horizon Discrete Time Control Problems for Bounded Processes.* Advances in difference equations, vol. 2008, article ID 654267.

[7] BERAN, Jakub. *Maximum Principle for Infinite Horizon Discrete Time Optimal Control Problems.* Diploma thesis, Univerzita Komenského v Bratislave, 2011.

[8] ARAUJO, A. & SCHEINKMAN, J. A. *Notes on comparative dynamics.* In: General Equilibrium, Growth,and Trade, J. R. Green and J. A. Scheinkman, Eds.,Academic Press, New York, NY, USA, 1979, pp. 217–226.

[9] MICHEL, P. *Some clarifications on the transversality condition.*In: Econometrica, vol. 58, no. 3, 1990, pp. 705–723,.

[10] STOKEY N. L. & LUCAS Jr. R. E. & PRESCOTT E. C. *Recursive Methods in Economic Dynamics.* Harvard University Press, Cambridge, Mass, USA, 1989.

[11] BRUNOVSKÝ, Pavol & HALICKÁ, Margaréta & JURČA, Pavol: *Optimálne riadenie.* Epos, 2009.

[12] YOSIDA, Kosaku. *Functional Analysis, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences).* vol. 123 (6th ed.), Berlin, New York: Springer-Verlag, 1980.

[13] PALMER, K. J. *Exponential Dichotomies, the Shadowing Lemma and Transversal Homoclinic Points.* In: Dynamics reported, Vol. 1, John Wiley & Sons Ltd. and B. G. Teubner, Stuttgart, 1988, pp 265–306.

[14] COPPEL, William. *Dichotomies in Stability Theory.* Springer-Verlag, New York, 1978.

[15] BOLTYANSKII, Vladimir Grigorevich. *Commande optimale des systèmes discrets.* MIR, Moscow, 1976.

[16] MCKENZIE, L. W. *Optimal Economic Growth, Turnpike Theorems and Comparative Dynamics. Handbook of Mathematical Economics.* Vol. III, Chap. 26, Elsevier, North-Holland, 1986.

[17] IOFFE, A. D. & TIHOMIROV,V. M. *Theory of Extremal Problems.* In: Studies in Mathematics and Its Applications,vol. 6, North-Holland, Amsterdam, The Netherlands, 1979.

[18] LE VAN, C. & SAGLAM,H. C. *Optimal Growth Models and the Langrange Multiplier.* In: Journal of Mathematical Economics, vol. 40, 2004, pp. 393-410.

[19] MICHEL, P. *Some Clarifications on the Transversality Condition.* In: Econometrica, 58, 1990, pp. 705–723.

[20] PELEG, B. & RYDER, Jr., H. E. *On Optimal Consumptions Plans in a Multi-Sector Economy.* In: Rev. Econ. Stud. 39, 1972, pp 159-170.

[21] DECHERT, W. D. *Lagrange Multipliers in Optimal Control Models.* Journal of Mathematical Economics, 9:285 302, 1982. Available: http://dx.doi.org/10.1016/0304-4068(82)90033-7.

[22] POLOQUIN, R. A. & ROCKAFELLAR, R. T. *A calculus of epi-derivatives applicable to optimization.* Canadian Journal of Mathematics 45 (4), 1993, pp. 879-896.

[23] ALIPRANTIS, Charalambos D. & BORDER, Kim C. *Infinite-Dimensional Analysis.* Springer, Berlin, Germany, 2nd edition, 1999.

[24] BRUNOVSKÝ, Pavol. *Nelineárna analýza.* Katedra aplikovanej matematiky a štatistiky, FMFI UK, 2006.

[25] RUDIN, Walter. *Functional Analysis.* McGraw-Hill Science/Engineering/Math, 1991. ISBN 978-0-07-054236-5.

[26] HABALA, Petr & HÁJEK, Petr & ZIZLER, Václav. *Introduction to Banach Spaces I.* Matfyzpress, 1996.

[27] DIESTEL, Joseph. *Sequences and Series in Banach Spaces*, Springer-Verlag, 1984. ISBN: 978-1-4612-9734-5.

[28] TESCHL, Gerald. *Topics in Real and Functional Analysis.* Fakultat fur Mathematik Nordbergstraße 15, Vienna, Austria February, 2011.

Appendix A

Basic Concepts in Functional Analysis

As we consider infinite horizon problems, we have to work with infinite sequences and their spaces. Therefore, an understanding of basic principles of functional analysis is neccesary for our research. In Appendix we summarize the basic concepts, principles and methods of functional analysis ussed in the thesis. We go though Banach spaces and operators on them, dual spaces. Most of the theory comes from the books [26], [27] and [28].

A.1 Metric and Banach Spaces

This section is devoted to metric, linear normed and Banach spaces. We give the definition of the spaces, illustrate them by several examples and describe some of their properties. We also introduce the space of all bounded sequences ℓ_∞^n and other ℓ_p^n spaces that are crucial for the thesis.

Definition A.1. Metric space

Let X be a nonempty set and $d : X \times X \to \mathbb{R}$ be a real function such that for all $x, y, z \in X$ one has

(i) $d(x, y) \geq 0$ and $d(x, y) = 0 \Leftrightarrow x = y$ (positivity)

(ii) $d(x, y) = d(y, x)$ (symmetry)

(iii) $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

The pair (X, d) is called a *metric space* and the function d a *metric*.

Example 2. Let ℓ_∞^n be the set of sequences $\mathbf{x} = \{x_t\}_{t=0}^\infty$, $x_t \in \mathbb{R}^n$ such that

$$\sup_{t \in \mathbb{N}_0} |x_t| < \infty$$

where $|\cdot|$ is a norm in the space $\mathbb{R}^n, n \in \mathbb{N}$. So it is a space of all bounded sequences. Denote

$$d(\mathbf{x}, \mathbf{y}) = \sup_{t \in \mathbb{N}_0} |x_t - y_t| \quad \text{for all } \mathbf{x}, \mathbf{y} \in \ell_\infty^n.$$

Then (ℓ_∞^n, d) is a metric space.

Example 3. For $p \in \langle 1, \infty)$ we introduce the set ℓ_p^n of sequences $\mathbf{x} = \{x_t\}_{t=0}^\infty$, $x_t \in \mathbb{R}^n$ such that

$$\sum_{t=0}^\infty |x_t|^p < \infty$$

where $|\cdot|$ is a norm in the space $\mathbb{R}^n, n \in \mathbb{N}$. Denote

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{t=0}^\infty |x_t - y_t|^p \right)^{\frac{1}{p}} \quad \text{for all } \mathbf{x}, \mathbf{y} \in \ell_p^n.$$

(ℓ_p^n, d) is also a metric space.

Definition A.2. Let (X, d) be a metric space. A *Cauchy sequence* is a sequence $\{x^{(n)}\}_{n=0}^\infty$, $x^{(n)} \in X$ for all $n \in \mathbb{N}_0$ such that for all $\varepsilon > 0$ there is $N_\varepsilon \in \mathbb{N}$ such that for all $n, m > N_\varepsilon$ $d(x^{(n)}, x^{(m)}) < \varepsilon$.

Definition A.3. A metric space is *complete* if all Cauchy sequences in this space converge.

Proposition A.1. If (X, d) is a metric space and $\{x^{(n)}\}_{n=0}^\infty$ is its Cauchy sequence, then $\{x^{(n)}\}_{n=0}^\infty$ is bounded in X, i.e. there exists $y \in X$ and $C \in \mathbb{R}$ such that $d(x^{(n)}, y) \leq C$ for all $x^{(n)} \in \{x^{(n)}\}_{n=0}^\infty$.

Definition A.4. Let (X, d) be a metric space. A *metric subspace* (Y, d_Y) of (X, d) consists of a subset $Y \subset X$ whose metric d_Y is the restriction of d to Y, that is $d_Y(x, y) = d(x, y)$ for all $x, y \in Y$.

Whenever we talk about a subspace Y of a metric space (X, d), we always consider it in terms of properties of the corresponding metric subspace (Y, d_Y).

Definition A.5. A subset Y of a metric space (X, d) is *closed* if it contains all its limit points, i.e. for all $\{x^{(n)}\}_{n=0}^{\infty}$, $x^{(n)} \in Y$ such that if $x^{(n)} \xrightarrow{n\to\infty} \bar{x}$, then $\bar{x} \in Y$.

Definition A.6. Normed linear space

Let $X \neq \emptyset$ be a vector space and $\|\cdot\| : X \to \mathbb{R}$ be a real function such that for all $x, y \in X$

(i) $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$

(ii) $\|\lambda x\| = |\lambda|\|x\|$ for every svalar λ

(iii) $\|x + y\| \leq \|x\| + \|y\|$.

The pair $(X, \|\cdot\|)$ is a *normed linear space* space and function $\|\cdot\|$ is a *norm.*

Proposition A.2. Let (X, d) be a normed linear space and define $d(x, y) = \|x - y\|$ for all $x, y \in X$. Then $(X, \|\cdot\|)$ is a metric space.

Example 4. The spaces ℓ_p^n, where $p \in \langle 1, \infty)$ and ℓ_∞^n are normed linear spaces, if we define the norms

$$\|\mathbf{x}\|_\infty = \sup_{t\in\mathbb{N}_0} |x_t|$$
$$\|\mathbf{x}\|_p = \left(\sum_{t=0}^{\infty} |x_t|^p\right)^{\frac{1}{p}} \qquad p \in \langle 1, \infty)$$

respectively.

Definition A.7. Banach space

Let $(X, \|\cdot\|)$ be a normed linear space. If the corresponding metric space (X, d) is complete we say $(X, \|\cdot\|)$ is a *Banach space*. (In the thesis, we left out the symbol $\|\cdot\|$ in the notation of the normed spaces).

Proposition A.3. The spaces ℓ_p^n, where $p \in \langle 1, \infty \rangle$ are Banach spaces.

Remark 3. In literature, spaces c and c_0 are often cited. $c \subset \ell_\infty$ is a space of convergent sequences and $c_0 \subset \ell_\infty$ is a space of sequences converging to 0. Both are Banach spaces.

A.2 Operators

Definition A.8. Let X, Y be normed linear spaces. By an operator T from X to Y we understand a map $T : X \supset \mathcal{D}(T) \to Y$. $\mathcal{D}(T)$ to be called *domain* of T. (In the thesis, we consider $\mathcal{D}(T) = X$ unless it is stated otherwise.)

Banach spaces and operators acting upon them form the basis of functional analysis. In this section we summarize necessary definitions and theorems associated with operators.

Definition A.9. The *range* $\mathcal{R}(T)$ of operator $T : X \to Y$ is the subset of Y of the values of T:

$$\mathcal{R}(T) = \{y \in Y, y = T(x) \text{ for some } x \in \mathcal{D}(T)\}.$$

Definition A.10. The *null space* $\mathcal{N}(T)$ of operator $T : X \to Y$

$$\mathcal{N}(T) = \{x \in \mathcal{D}(T), T(x) = 0\}.$$

Example 5. Let X be a linear space. The *identity* operator $I : X \to X$ is defined by

$$I(x) = x \quad \text{for all } x \in X.$$

Addition and scalar multiplication of operators are defined similarly to that of standard functions.

Definition A.11. Let T_1 and T_2 be mappings from normed space X to normed space Y. We define operator $T_1 + T_2 : X \to Y$ with the domain $\mathcal{D}(T_1 + T_2) = \mathcal{D}(T_1) \cap \mathcal{D}(T_2)$ and the rule

$$(T_1 + T_2)(x) = T_1(x) + T_2(x) \quad \text{for all } x \in \mathcal{D}(T_1 + T_2).$$

Let $\lambda \in \mathbb{R}$. We define operator $\lambda T_1 : X \to Y$ with the domain $\mathcal{D}(T_1)$ and the rule

$$(\lambda T_1)(x) = \lambda T_1(x) \quad \text{for all } x \in \mathcal{D}(T_1).$$

Definition A.12. Let X, Y be normed spaces and $T : X \to Y$ be a map between them. T is called *linear* or *homomorphism* if

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y),$$

for all $x, y \in X$ and $\alpha, \beta \in \mathbb{R}$.

Example 6. The operator $\sigma : X \to X$, where X is Banach space, defined by $\sigma(x_0, x_1, x_2, ...) = (x_1, x_2, ...)$ is called *shift* operator (sometimes it is called left shift operator as the shift is taken from the right to the left). For finite vector $x \in X$ we define $\sigma(x_0, x_1, x_2, ..., x_n) = \sigma(x_1, x_2, ..., x_n, 0)$. Shift operator is linear as

$$\sigma(\alpha(x_0, x_1, x_2, ...) + \beta(y_0, y_1, y_2, ...)) = \sigma((\alpha x_0, \alpha x_1, \alpha x_2, ...) + (\beta y_0, \beta y_1, \beta y_2, ...)) =$$
$$(\alpha x_1, \alpha x_2, ...) + (\beta y_1, \beta y_2, ...) = \alpha(x_1, x_2, ...) + \beta(y_1, y_2, ...) =$$
$$\alpha\sigma(x_0, x_1, x_2, ...) + \beta\sigma(y_0, y_1, y_2, ...).$$

Definition A.13. Let X, Y be normed spaces and $T : X \to Y$ be a map between them. T is called *continuous* in $x \in X$, if

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \|x - y\|_X < \delta \Rightarrow \|T(x) - T(y)\|_Y < \varepsilon$$

T is said to be continuous if it is continuous over its domain $\mathcal{D}(T)$.

Definition A.14. Let X, Y be normed spaces and $T : X \to Y$ be a map between them. T is called *bounded* if there exists a constant $C > 0$ such that

$$\|T(x)\|_Y \leq C\|x\|_X \quad \text{for all } x \in X.$$

Definition A.15. Let X, Y be normed spaces and $T : X \to Y$ be a map between them. T is called *closed* if every sequence $\{x^{(n)}\}_{n=0}^{\infty} \in X$ converging to $x \in X$ holds

$$\lim_{n\to\infty} T(x^{(n)}) = T(x).$$

Proposition A.4. Let X, Y be normed spaces and T be a linear operator between them. Then T is continuous if and only if it is bounded.

Definition A.16. Let T be a bounded linear operator from a normed space X to a normed space Y. The *norm* of T is defined as

$$\|T\| = \sup_{\|x\|_X=1} \|T(x)\|_Y.$$

We denote the linear space of all bounded operators from X to Y with the norm $\|\cdot\|$ by $\mathcal{B}(X, Y)$.

Remark 4. The norm of an operator satisfies the properties of norm defined in Definition A.6.

Remark 5. The operator norm is the smallest C from the definition of boundness (Definition A.14), i.e.

$$\|T(x)\|_Y \leq \|T\|\|x\|_X \quad \text{for all } x \in X.$$

Therefore, in order to prove boundness of an operator T, it is sufficient to prove that its norm is finite.

Definition A.17. Let X, Y be Banach spaces and $T : X \to Y$. Then T is called

1. *injective* if $T(x) = T(y)$ implies $x = y$,

2. *surjective* if $\mathcal{R}(T) = Y$,

3. a *bijection* if it is both injective and surjective,

4. an *isomorphism* if it is a bijective homomorphism.

A.3 Dual Spaces

Definition A.18. Let X be a normed space. A linear operator $x^* : X \to \mathbb{R}$ is called *linear functional* and we define $< x^*, x >= x^*(x)$ and $\|x^*\| = \sup_{\|x\|=1} x^*(x)$. The space of all continuous linear functionals from X to $\mathbb{R}$ is called *dual space* of X and is denoted by X^*.

To distinguish X from X^* we occasionally call the former primal space. This section describes the concept of dual spaces on Banach spaces and identify dual spaces for ℓ_p^n, $p \in \langle 1, \infty \rangle$.

Proposition A.5. The dual space of a normed space X is a Banach space.

Definition A.19. Let X, Y be Banach spaces and $T \in \mathcal{B}(X, Y)$. We define the *dual operator* or *adjoint operator* $T^* \in \mathcal{B}(Y^*, X^*)$ for $y^* \in Y^*$ by

$$T^*(y^*)(x) = y^*(T(x)) \quad \text{for all } x \in X.$$

Definition A.20. A Banach space X is *reflexive*, if $(X^*)^* = X$.

Proposition A.6. For $p \in (1, \infty)$, the dual space of ℓ_p is ℓ_q with q such that $\frac{1}{q} = 1 - \frac{1}{p}$.

Proposition A.7. The dual space of ℓ_1 is ℓ_∞.

Proposition A.8. $(\ell_\infty)^* = \ell_1 \oplus \ell_s$, i.e. ℓ_∞, ℓ_1 are not reflexive.

A.4 Convergence

Closed and bounded sets of the infinite dimensional normed linear spaces are not necessarily sequentially compact, i.e. a bounded sequence may not contain a convergent subsequence. However, convergence can frequently be replaced by a weaker concept.

Definition A.21. Let X be a normed linear space and sequence $\{x^{(n)}\}_{n=1}^{\infty}$ such that $x^{(n)} \in X$.

1. We say that $\{x^{(n)}\}_{n=0}^{\infty}$ *converges strongly* to x, $x^{(n)} \to x$, if

$$\lim_{n\to\infty} \|x^{(n)} - x\| = 0.$$

2. We say that $\{x^{(n)}\}_{n=0}^{\infty}$ *converges weakly* to x, $x^{(n)} \xrightarrow{w} x$, if

$$\lim_{n\to\infty} \langle x^*, x^{(n)} \rangle = \langle x^*, x \rangle \quad \text{for all } x^* \in X^*.$$

3. We say that $\{x^{*(n)}\}_{n=0}^{\infty} \in X^*$ *converges weakly** to x^*, $x^{*(n)} \xrightarrow{w*} x^*$, if

$$\lim_{n\to\infty} \langle x^{*(n)}, x \rangle = \langle x^*, x \rangle \quad \text{for all } x \in X.$$

Remark 6. Suppose that X is a normed space and the sequence $\{x^{(n)}\}_{n=0}^{\infty}$, $x^{(n)} \in X$ converges strongly. Then it also converges weakly, i.e. if $x^{(n)} \to x$, then $x^{(n)} \xrightarrow{w} x$. It is because we have for all $x^* \in X^*$, $|\langle x^*, x^{(n)} \rangle - \langle x^*, x \rangle| \leq \|x^*\| \|x^{(n)} - x\|$. Since $\|x^{(n)} - x\| \to 0$, $|\langle x^*, x^{(n)} \rangle - \langle x^*, x \rangle| \to 0$, as well.

Appendix B

Important inequalities

In the thesis we applied several inequalities. We summarize them in this chapter.

Proposition B.1. (The Hölder inequality)
Let $p, q \in \langle 1, \infty \rangle$ be such that

$$\frac{1}{p} + \frac{1}{q} = 1$$

where we define $\frac{1}{\infty} = 0$ and let $\mathbf{x} \in \ell_p$, $\mathbf{y} \in \ell_q$. Then $\mathbf{xy} \in \ell_1$ and

$$\|\mathbf{x}, \mathbf{y}\|_1 \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

Proposition B.2. (The Minkowski inequality)
Let $p \in \langle 1, \infty \rangle$ and $\mathbf{x}, \mathbf{y} \in \ell_p$ then

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

Proposition B.3. (The Jensen's inequality) Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function, x_t be in its domain and a_t be positive weights for all $t \in \mathbb{N}_0$. Then

$$f\left(\frac{\sum_{t=0}^{\infty} a_t x_t}{\sum_{t=0}^{\infty} a_t}\right) \leq \frac{\sum_{t=0}^{\infty} a_t f(x_t)}{\sum_{t=0}^{\infty} a_t}.$$