FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

BACHELOR THESIS

Bohdan Šmaňko

Well-posed optimization problems

Department of Probability and Mathematical Statistics

Supervisor of the bachelor thesis: doc. RNDr. Petr Lachout, CSc.
Study programme: General Mathematics

Prague 2025

I declare that I carried out this bachelor thesis on my own, and only with the cited sources, literature and other professional sources. I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

Dedicated to my infinitely patient mom.

Title: Well-posed optimization problems

Author: Bohdan Šmaňko

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Petr Lachout, CSc., Department of Probability and Mathematical Statistics

Abstract: We study the relations between different notions of well-posed optimization problems, with focus on the well-posedness of minimization problems in metric spaces. We begin by introducing the concepts of Tikhonov, Levitin-Polyak, strong and Hadamard well-posedness, providing corresponding characterizations along with illustrative numerical examples. The main result of this thesis is the complete proof of equivalences between these notions of well-posedness under specific assumptions. Then we also examine well-posedness in a generalized setting. As an application, we demonstrate that the problem of minimizing the variance of a two-asset portfolio is well-posed in every sense. Next, we explore the notion of ill-posed problems and present Tikhonov regularization method as an approach to solve them. Lastly, we discuss the practical applications of well-posedness across various fields.

Keywords: well-posed optimization problem, ill-posed optimization problem, problem parameters, functional dependence on parameters

Název práce: Well-posed optimalizační úlohy

Autor: Bohdan Šmaňko

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Petr Lachout, CSc., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Zabýváme se vztahy mezi různými druhy well-posed optimalizačních úloh, se zaměřením na well-posed minimalizační úlohy v metrických prostorech. Začínáme představením konceptů Tikhonovovy, Levitin-Polyakovy, silné a Hadamardovy well-posedness, kde poskytujeme odpovídající charakterizace doplněné ilustrativními numerickými příklady. Hlavním výsledkem této práce je kompletní důkaz ekvivalencí mezi zmíněnými druhy well-posedness za konkrétních předpokladů. Pak také zkoumáme well-posedness v obecnějším smyslu. V praktickém příkladu ukazujeme, že minimalizace rozptylu v portfoliu dvou aktiv je well-posed v každém smyslu. Dále se zabýváme pojmem ill-posed úloh a uvádíme Tikhonovovu regularizační metodu jako možný postup k jejich řešení. Na závěr probíráme praktická využití well-posedness v různých oborech.

Klíčová slova: well-posed optimalizační úloha, ill-posed optimalizační úloha, parametry úlohy, funkční závislost na parametrech

Contents

Introduction

Throughout all mathematical disciplines, we encounter countless problems with the goal of finding a desired solution. From a fundamental standpoint it is reasonable to ask whether it makes sense to attempt solving such problems without first establishing that the sought solution actually exists. In the early 20th century, the French mathematician Jacques Hadamard introduced the notion of well-posedness for a problem, formulating the concept of such questions. According to Hadamard, a problem is well-posed if it satisfies the following three key conditions. Specifically, a well-posed problem is one for which a solution exists, the solution is unique, and is continously dependent on the data upon which the problem is based, meaning that small changes in the input data lead to correspondingly small changes in the solution. At the start of 1960s, the Russian mathematician Andrey Tikhonov introduced another notion of well-posedness. While his definition also requires existence and uniqueness of the solution, he further imposes the convergence of every minimizing sequence to a unique minimum point. In other words, sequences of approximate solutions must converge to the optimal solution. Soon after, Levitin and Polyak introduced a generalization of Tikhonov's concept. Approximately 3 decades later, Beer and Lucchetti presented a further generalized concept of well-posedness.

This thesis will explore and analyze each of these concepts, comparing them and charactarizing their relationships. Additionaly present appropriate numerical examples to highlight the differences in the notions of well-posedness. We will see that under specific assumptions, the concepts coincide, which we will fully prove. Then we will present our next contribution, where we will prove that the problem of the minimizing variance of a two-asset portfolio is well-posed in compliance with all aforementioned notions of well-posedness. To support this, we will also show a simple example with real data of stock options for said problem. The introduction of well-posed optimization problems conversely implies the notion of optimization problems that are not well-posed, ill-posed problems. Which we will address in the fourth chapter and present an approach to approximately solve such problems. The final chapter involves a few examples of applied well-posedness in various fields.

1 Well-posedness

1.1 Problem introduction

Our main focus in the first chapter is to introduce and characterize various notions of well-posedness for optimization problems involving the minimization of a function f over a metric space X. Well-posedness in more general topological spaces shares similar ideas, but the results are broader, while the concepts are significantly more complex. Therefore, we choose to work with metric spaces. In this section, we establish the necessary definitions to describe our problem.

Let X be a metric space with a metric $d : X \times X \to \mathbb{R}$ and we clarify the following notations, which we will use throughout the paper. For a subset $A \subseteq X$ and a point $x \in X$ we denote the metric d as $d(x, A) = \inf\{d(x, a) : a \in A\}$. Sequences $\{x_n\}_{n=1}^{\infty}$ are denoted as simply $\{x_n\}$ unless explicitly mentioned otherwise. Limits and convergences expressed with $\to$ are meant as the corresponding index goes to ∞. For example $x_n \to x_0$ denotes $\lim_{n\to\infty} x_n = x_0$ unless specified differently.

With the following definitions, we will construct a set of functions suitable for our problem.

Definition 1.1.1. Let $f : X \to \overline{\mathbb{R}}$ be an extended real-valued function. Then f is said to be proper if it never assumes the value $-\infty$ and it is not identically ∞.

Definition 1.1.2. Let $f : X \to \overline{\mathbb{R}}$, then f is bounded from below if there exists a real number K such that $\forall x \in X : f(x) \geq K$.

Definition 1.1.3. Given a metric space X, let $\mathcal{F}(X)$ be a set of all extended real-valued functions on X that satisfy the following properties:

$$\mathcal{F}(X) := \{f : X \to \overline{\mathbb{R}} : f \text{ is proper and bounded from below}\}.$$

$\mathcal{F}(X)$ is equipped with the uniform metric:

$$e(f_1, f_2) := \sup_{x\in X} \left\{ \tfrac{|f_1(x)-f_2(x)|}{1+|f_1(x)-f_2(x)|} \right\}, \quad f_1, f_2 \in \mathcal{F}(X).$$

Next, we define the optimization problem that we will work with moving forward.

Definition 1.1.4. Let X be a metric space and $f \in \mathcal{F}(X)$. Consider the problem determined by the pair (X, f), which we will denote as $\min(X, f)$:

$$\min_{x\in X} f(x)$$

that consists of finding $x_0 \in X$ such that

$$f(x_0) = \inf_{x\in X} f(x) := \inf(X, f).$$

The set of optimal solutions for the problem $\min(X, f)$ is denoted by $\operatorname{argmin}(X, f)$. If a solution x_0 exists and is unique then:

$$\{x_0\} = \operatorname{argmin}(X, f).$$

The rest of the chapter will concentrate on the well-posedness of $\min(X, f)$. We are interested in the well-posedness of a minimization problem, but the ensuing concepts can be applied to a maximization problem under analogous conditions due to the relationship between minimization and maximization, which is given by $\max(f) = -\min(-f)$. The classical idea of well-posedness requires existence of a unique solution and its continous dependancy on the data of the considered problem.

1.2 Tikhonov well-posedness

We present the concept of Tikhonov well-posedness introduced in 1966 by A.N. Tikhonov, which characterizes the classical idea with minimizing sequences.

Definition 1.2.1. Let X be a metric space and let $f \in \mathcal{F}(X)$. A sequence $\{x_n\} \in X$ is called *Tikhonov minimizing* (briefly T. minimizing) of $\min(X, f)$ if

$$f(x_n) \to \inf_{x \in X} f(x), \quad n \to \infty.$$

Definition 1.2.2. A minimization problem $\min(X, f)$ is *Tikhonov well-posed* (briefly T.w.p.) if all following conditions hold:

(i) there exists a unique solution $x_0 \in X$, i.e., $\operatorname{argmin}(X, f)$ is nonempty and a singleton;

(ii) every T. minimizing sequence $\{x_n\}$ of $\min(X, f)$ converges to x_0.

The concept of Tikhonov well-posedness connects the theory of optimization with its practical use in numerical analysis. We can view the T. minimizing sequence as a sequence of approximate solutions for $\min(X, f)$ and Tikhonov well-posedness ensures that every such sequence converges to the optimal solution of $\min(X, f)$. Next we show a simple example of applied T.w.p. inspired by Ferrentino, Boniello [1].

Example 1.2.1. Let $X = \mathbb{R}$ and $f(x) = x^2 e^{-x}$, in that case the unique solution of the minimization problem $\min(X, f)$ is clearly $x_0 = 0$, where $f(x_0) = 0 = \inf_{\mathbb{R}} f$. Assume $\{x_n\} = \{n\}_{n \in \mathbb{N}}$, then $\{x_n\}$ is T. minimizing for $\min(X, f)$, because $f(x_n)$ is decreasing for $n > 2$ and converges to $0 = \inf_{\mathbb{R}} f$ as $n \to \infty$. But $x_n = n \to \infty$, therefore the T. minimizing sequence does not converge to the optimal solution and $\min(X, f)$ is not T.w.p.
If we modify the problem and put $f(x) = x^2$, then we can see that $\inf_X f = 0$, where the unique solution is $x_0 = 0$. Now let $\{x_n\}$ be a T. minimizing sequence in X, which yields $f(x_n) \to f(x_0) = 0$. In other words $x_n^2 \to 0$ and due to Continous mapping theorem we get that $x_n \to x_0 = 0$, therefore $\min(X, f)$ is T.w.p.

Since our goal is to find the infimum of f over a metric space X, then lower semicontinuity of f is desirable as it plays a crucial role in the existence of a solution for $\min(X, f)$.

Definition 1.2.3. A function $f \in \mathcal{F}(X)$ is said to be *lower semicontinous* (briefly l.s.c.) at $x_0 \in X$ if $\forall \varepsilon > 0$, there exists $\delta > 0$ such that

$$f(x_0) - \varepsilon < f(x) : \forall x \text{ that satisfy } d(x, x_0) < \delta$$

f is called *lower semicontinous* if f is l.s.c. at every point of its domain.
For our purposes the following equivalence might be more telling:

$$f \text{ is l.s.c. at } x_0 \Leftrightarrow \forall \{x_n\} \in X, x_n \to x_0 : \liminf_{n \to \infty} f(x_n) \geq f(x_0).$$

Tikhonov well-posedness requires that a sequence of approximate solutions approaches the point where f attains its infimum. Therefore, it is intuitive to charactarize T.w.p. for f l.s.c. using level sets. This approach, introduced by Furi and Vignoli, is known as the Furi-Vignoli criterion.

Definition 1.2.4. Let $f : X \to \overline{\mathbb{R}}$. We define its *level set* at height $a \in \mathbb{R}$ as

$$f^a := \{x \in X : f(x) \leq a\}.$$

We will use the formulation of the Furi-Vignoli criterion as presented in Proposition 10.1.6, Lucchetti [2]. To do so, we will need the definitions of a complete metric space and Cauchy sequences. Additionaly we introduce the following Lemma 1.2.1. about the properties of level sets for f l.s.c., which will be helpful in the proof of the criterion.

Lemma 1.2.1. *Let $f \in \mathcal{F}(X)$ be a lower semicontinous function and f^a a level set for $a > \inf_X f$, then f^a is nonempty and closed.*

Proof. The function f is proper and bounded from below, therefore $\inf_X f$ exists and is not equal to $-\infty$. Since f is l.s.c., then f attains its infimum at some $x_0 \in X$, and we get $f(x_0) = \inf_X f < a$. Thus $x_0 \in f^a$ and f^a is nonempty.

Let $\{y_n\} \in f^a$ be a sequence that converges to some $y_0 \in X$, then from the definition of l.s.c. we get

$$f(y_0) \leq \liminf f(y_n) \leq a.$$

Therefore $f(y_0) \leq a$, which yields $y_0 \in f^a$. Every convergent sequence in f^a has its limit in f^a, hence f^a is a closed subset of X.

□

Definition 1.2.5. A sequence $\{x_n\} \in X$ is called a *Cauchy* sequence if

$$\forall \varepsilon > 0 \text{ there exists } n_0 \in \mathbb{N} \text{ such that } \forall m, n > n_0 : d(x_m, x_n) < \varepsilon.$$

Definition 1.2.6. A metric space X is called *complete* if every Cauchy sequence in X converges to a point also in X.

Theorem 1.2.1. ***(Furi-Vignoli criterion)***
Let X be a complete metric space and let $f \in \mathcal{F}(X)$ be lower semicontinous, then the following assertions are equivalent:

(i) $\min(X, f)$ *is Tikhonov well-posed;*

(ii) $\inf_{a > \inf_X f} \operatorname{diam} f^a = 0,$

where $\operatorname{diam} f^a := \sup_{x,y \in f^a} d(x, y).$

Proof.
(i) $\Rightarrow$ **(ii)**; Assume $\min(X, f)$ is T.w.p. and (ii) does not hold, which means that $\operatorname{diam} f^a > 0$ for $a = \inf_X f + \varepsilon, \varepsilon > 0$. From Lemma 1.2.1., we get that f^a is nonempty and closed. Therefore there exist points $x_0, y_0 \in f^a$ such that $d(x_0, y_0) > 0$. Then we can find $\delta > 0$ and two T. minimizing sequences $\{x_n\}$ and $\{y_n\}$ such that $\forall n \in \mathbb{N} : d(x_n, y_n) \geq \delta$, where $\{x_n\}$ converges to x_0 and $\{y_n\}$ converges to y_0, while $f(x_0), f(y_0) \leq \inf_X f + \varepsilon$. This implies that $\operatorname{argmin}(X, f)$ is not a singleton and therefore (i) is not true, hence (i) $\Rightarrow$ (ii) by contradiction.

(ii) $\Rightarrow$ **(i)**; The (ii) implies that there exists $a > \inf_X f$ such that $\operatorname{diam} f^a = 0$. For said a we get a level set f^a, where $\forall x, y \in f^a : x = y$, because $\sup_{x,y \in f^a} d(x, y) = 0$ and as a result f^a is a singleton, let us denote that point as x_0. From the definition of a level set we get that x_0 is the only point in X for which $f(x_0) \leq a$ is true and since $a > \inf_X f$, then $f(x_0) = \inf_X f$. Consequently x_0 is a unique minimum point for f in X. Let $\{x_n\}$ be a T. minimizing sequence, then the T. minimizing property $f(x_n) \to \inf_X f = f(x_0)$ and (ii), which provides that points of $\{x_n\}$ cannot be too far apart as we approach the minimum point, which implies that $\{x_n\}$ is a Cauchy sequence. Since X is a complete space, then the Cauchy sequence $\{x_n\}$ converges to a point in X. Again $\{x_n\}$ is a T. minimizing sequence of $\min(X, f)$ and so $f(x_n) \to f(x_0)$, hence the point of convergence of $\{x_n\}$ is the minimum point, which is unique as aforementioned. We have proved that every T. minimizing sequence converges to a unique minimum point, and thus $\min(X, f)$ is T.w.p.

□

Another characterization of Tikhonov well-posedness involves estimating the difference between the value of an approximate solution $f(x)$ and the value of the optimal solution $f(x_0)$ in terms of $d(x, x_0)$. Such concept and its implications are explored in Chapter 1, Section 2, Dontchev, Zolezzi [3] and Section 10.1, Lucchetti [2]. Here we will only showcase how $f(x) - f(x_0)$ and $d(x, x_0)$ relates to Tikhonov well-posedness by utilizing the notion of a forcing function.

Definition 1.2.7. Let $T \subset [0, \infty)$ and a function $c : T \to [0, \infty)$, c is said to be a forcing function if the following conditions hold:

(i) $0 \in T, c(0) = 0;$

(ii) for a sequence $\{t_n\} \in T : c(t_n) \to 0 \Rightarrow t_n \to 0, \quad n \to \infty.$

Theorem 1.2.2. *Let X be a metric space and $f \in \mathcal{F}(X)$. For the given minimization problem* $\min(X, f)$ *the following are equivalent:*

(i) $\min(X, f)$ *is Tikhonov well-posed;*

(ii) there exists a forcing function c and a point $x_0 \in X$ such that

$$f(x) \geq f(x_0) + c[d(x, x_0)], \forall x \in X.$$

Proof.
(ii) ⇒ (i); Assume there exists a forcing function c and $x_0 \in X$ in compliance with (ii). Let $\{x_n\}$ be a T. minimizing sequence of $\min(X, f)$. According to the T. minimizing property of $\{x_n\}$ and assumption that (ii) holds, we get $f(x_n) \to f(x_0)$, then $c[d(x_n, x_0)] \to 0$, which implies $d(x_n, x_0) \to 0$ as per the definition of a forcing function. We have proved that every T. minimizing sequence of $\min(X, f)$ converges to a single point x_0, which means that $\min(X, f)$ is T.w.p. with x_0 as its solution.

(i) ⇒ (ii); Let $\min(X, f)$ be T.w.p. with a unique solution $x_0 \in X$. We define $c(t)$ for $t \geq 0$ as

$$c(t) = \inf_{x \in X}\{f(x) - f(x_0) : d(x, x_0) = t\}.$$

We will show that $c(t)$ is a forcing function. For $t = 0$ we get $c(0) = f(x_0) - f(x_0) = 0$ and since x_0 is the solution of $\min(X, f)$, then $f(x) - f(x_0) \geq 0$ for all $x \in X$, which means $c(t) \geq 0, \forall t \geq 0$.
Let $\{t_n\}$ be a sequence where $t_n \geq 0, \forall n \in \mathbb{N}$ such that $c(t_n) \to 0$, then there exists a sequence $\{y_n\}$ in X such that $f(y_n) \to f(x_0)$ where $d(y_n, x_0) = t_n, \forall n \in \mathbb{N}$. This means $\{y_n\}$ is T. minimizing and therefore $t_n \to 0$, which proves that $c(t)$ is a forcing function and so (i) ⇒ (ii).

□

In most research on the theory of well-posedness, there is a strong focus on the density of the set of all well-posed problems in a given problem product space. This leads to compelling results, demonstrating that in some sense the "majority" of the optimization problems is well-posed under specific conditions. We mention an example for Tikhonov well-posedness from Revalski [4].

Definition 1.2.8. Let X be a metric space, then we define $\mathcal{H}(X)$ as the set of all nonempty closed subsets of X:

$$\mathcal{H}(X) := \{A \subset X : A \neq \emptyset \text{ and } A \text{ is closed}\}.$$

$\mathcal{H}(X)$ is endowed with the Hausdorff distance ρ_H where:

$$\rho_H(A_1, A_2) := \max\left\{\sup_{x \in A_2} d(x, A_1), \sup_{x \in A_1}\{d(x, A_2)\}\right\}, \quad A_1, A_2 \in \mathcal{H}(X).$$

Theorem 1.2.3. *Let X be a complete metric space and $\mathcal{F}_l(X)$ be a set of all functions $f \in \mathcal{F}(X)$ that are lower semicontinous. Denote $\mathcal{P}_T \subset \mathcal{H}(X) \times \mathcal{F}_l(X)$ as a set of all Tikhonov well-posed problems, then $\mathcal{P}_T$ is dense in $\mathcal{H}(X) \times \mathcal{F}_l(X)$.*

Proof. Refer to Corollary 1.5., Revalski [4]

□

1.3 Levitin-Polyak well-posedness

In the previous section, we were concerned with minimizing f over the entire metric space X. In other words the problem $\min(X, f)$ was an unconstrained problem. However, in practice, we often encounter optimization problems involving boundaries or restrictions. Therefore we will now focus on the constrained problem $\min(A, f)$, where A is a subset of X. This provides us with the possibility to approximate the optimal solution using sequences, which do not necessarily lie within A. This concept, introduced by Levitin and Polyak, presents a strengthened notion of minimizing sequences and a new type of well-posedness, Levitin-Polyak well-posedness.

Definition 1.3.1. Let $A \subset X, A \neq \emptyset$ and $f \in \mathcal{F}(X)$. A sequence $\{x_n\} \in X$ is called *Levitin-Polyak minimizing* (briefly L.P. minimizing) of $\min(A, f)$ if it satisfies the following conditions:

(i) $f(x_n) \to \inf_{x \in A} f(x), \quad n \to \infty;$

(ii) $d\,(x_n, A) \to 0, \quad n \to \infty.$

Definition 1.3.2. A minimization problem $\min(A, f)$ is *Levitin-Polyak well-posed* (briefly L.P.w.p.) if all following conditions hold:

(i) there exists a unique solution $x_0 \in A$, i.e., $\operatorname{argmin}(A, f)$ is nonempty and a singleton;

(ii) every Levitin-Polyak minimizing sequence $\{x_n\}$ of $\min(A, f)$ converges to x_0.

To illustrate the difference between the notions of Tikhonov and Levitin-Polyak well-posedness, we solve the Exercise 10.1.14 from Lucchetti [2] in the following example.

Example 1.3.1. Consider the minimization problem $\min(A, f)$, where $X = \mathbb{R}^2$, $A := \{(x, y) \in X : y = 0\}$ and $f(x, y) = x^2 - x^4y^2$. Then for $(x, y) \in A$ we have $f(x, 0) = x^2$ and we see that $\min(A, f)$ is T.w.p. as shown in Example 1.2.1.
According to the result in the aforementioned example, we have the unique solution $(0, 0)$, where $f(0, 0) = \inf_A f = 0$. Now let $\{x_n, y_n\} = \{n, 1/n\}_{n \in \mathbb{N}}$ be a sequence in X and we get

$$f(x_n, y_n) = n^2 - \tfrac{n^4}{n^2} = 0, \quad \forall n \in \mathbb{N},$$

$$(n, \tfrac{1}{n}) \to (\infty, 0) \in A \Rightarrow d((x_n, y_n), A) \to 0, \quad n \to \infty.$$

As a result $\{x_n, y_n\}$ is L.P. minimizing sequence, but (x_n, y_n) does not converge to the solution $(0, 0)$, therefore $\min(A, f)$ is not L.P.w.p

Recalling the Furi-Vignoli criterion from the previous section, we can similarly characterize the Levitin-Polyak well-posedness using generalized level sets. We will present the version of this concept as shown in Theorem 30, Dontchev, Zolezzi [3].

Definition 1.3.3. Let X be a metric space and $f \in \mathcal{F}(X)$. Consider a subset $A \subset X, A \neq \emptyset$. We define a *generalized level set* L^a for the subset A and $a > 0$ as

$$L^a = \{x \in X : d(x, A) \leq a \text{ and } f(x) \leq \inf_A f + a\}.$$

Theorem 1.3.1. *Let X be a complete metric space with a closed subset $A \subset X, A \neq \emptyset$ and $f \in \mathcal{F}(X)$ is lower semicontinous then*

$$\operatorname{diam} L^a \to 0 \text{ as } a \to 0$$

implies that a minimization problem $\min(A, f)$ *is Levitin-Polyak well-posed.*

Proof. Let $\{x_n\}$ be any L.P. minimizing sequence of $\min(A, f)$, i.e. $d(x_n, A) \to 0$ and $f(x_n) \to \inf_A f$. Assume that $\operatorname{diam} L^a \to 0$ as $a \to 0$, then given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\operatorname{diam} L^a < \varepsilon \text{ for } 0 < a < \delta.$$

Then for every $a \in (0, \delta)$, there exists a sufficiently large $n_0 \in \mathbb{N}$ such that $f(x_k) \leq \inf_A f + a$ and $d(x_k, A) < a$ for all $k \geq n_0$, and thus $x_k \in L^a$, which implies that $\{x_n\}$ is a Cauchy sequence, because $\forall n, m \geq n_0 : x_n, x_m \in L^a$, therefore $d(x_n, x_m) < \varepsilon$. As a result $\{x_n\}$ converges to some $x_0 \in A$, because $d(x_n, A) \to 0$, where A is a closed subset of X, which is a complete metric space. Now using the result $x_n \to x_0$ and the definition of lower semicontinuity of f, we get $\liminf f(x_n) \geq f(x_0)$ and since $f(x_n) \to \inf_A f$, then $\inf_A f \geq f(x_0)$. By the definition of infimum we also have $\inf_A f \leq f(x_0)$ since $x_0 \in A$, and thus $f(x_0) = \inf_A f$. To show that x_0 is unique, let $a = 0$, then $x_0 \in L^a$, because $d(x_0, A) = 0$ and $f(x_0) = \inf_A f$. Assuming $\operatorname{diam} L^a \to 0$ as $a \to 0$, we get that x_0 is the only point in L^a for $a = 0$, therefore x_0 is a unique minimum point. We have shown that every L.P. minimizing sequence converges to $x_0 \in A$, where x_0 is a unique minimum point, therefore $\min(A, f)$ is L.P.w.p.

□

1.4 Strong well-posedness

Eventually, Beer and Lucchetti introduced a generalization of the Levitin-Polyak minimizing sequences, leading to the concept of strong well-posedness, a more robust version of Levitin-Polyak well-posedness.

Definition 1.4.1. Let $A \subset X, A \neq \emptyset$ and $f \in \mathcal{F}(X)$. A sequence $\{x_n\} \in X$ is called *strongly minimizing* of $\min(A, f)$ if it satisfies the following conditions:

(i) $\limsup f(x_n) \leq \inf_{x \in A} f(x), \quad n \to \infty;$

(ii) $d(x_n, A) \to 0, \quad n \to \infty.$

Definition 1.4.2. A minimization problem $\min(A, f)$ is *strongly well-posed* (briefly s.w.p.) if all following conditions hold:

(i) there exists a unique solution $x_0 \in A$, i.e., argmin (A, f) is nonempty and a singleton;

(ii) every strongly minimizing sequence $\{x_n\}$ of $\min(A, f)$ converges to x_0.

Although, as we will demonstrate later in Theorem 1.6.1, the two notions coincide in certain situations. But f being lower semicontinous is not enough as will be shown in the following example taken from p.151, Revalski, Zhivkov [5].

Example 1.4.1. Let $X = \mathbb{R}^2$, $A := \{(x, y) \in X : y = 0\}$ and consider the minimization problem $\min(A, f)$ for

$$f(x,y) = \begin{cases} x^2 + y^2, & (x,y) \in \left\{(x,y) \in X : (x,y) \neq (n, \frac{1}{n})_{n\in\mathbb{N}}\right\}, \\ -1, & (x,y) = (n, \frac{1}{n})_{n\in\mathbb{N}} \end{cases} \tag{1.1}$$

Where f is lower semicontinous, which can be shown thanks to the alternate definition of l.s.c. It is obvious that f is l.s.c. at $(x, y) \neq (n, \frac{1}{n})_{n\in\mathbb{N}}$, since $x^2 + y^2$ is a continous function on X. Now consider any sequence $\{a_{n_k}, b_{n_k}\}$ in X such that (a_{n_k}, b_{n_k}) converges to $(n, 1/n)$ for all $n \in \mathbb{N}$ as $k \to \infty$, then

$$f(a_{n_k}, b_{n_k}) = a_{n_k}^2 + b_{n_k}^2,$$

and therefore

$$\liminf_{k\to\infty} f(a_{n_k}, b_{n_k}) \geq 0 \geq -1 = f(n, \tfrac{1}{n}), \quad \forall n \in \mathbb{N},$$

which gives us that f is l.s.c. at all points in X. To show that $\min(A, f)$ is L.P. well-posed, let us first restrict f on A to find the optimal solution. Because $\forall (x, y) \in A : (x, y) \neq (n, \frac{1}{n})_{n\in\mathbb{N}}$ we get

$$f(x, y) = x^2, \quad \forall (x, y) \in A.$$

Which gives us the unique solution $(0, 0)$ and $f(0, 0) = \inf_A f = 0$ again as shown in Example 1.2.1.
Let $\{x_n, y_n\}$ be any L.P. minimizing sequence, then

$$x_n^2 + y_n^2 \to 0 \text{ and } d((x_n, y_n), A) \to 0, \text{ as } n \to \infty.$$

Where $d((x_n, y_n), A) \to 0$ gives us that $y_n \to 0$, consequently $x_n^2 \to 0 \Rightarrow x_n \to 0$ for $x_n \in \mathbb{R}$ due to Continous mapping theorem. Therefore $(x_n, y_n) \to (0, 0)$, hence $\min(A, f)$ is L.P.w.p. We will use a similar argument as in Example 1.3.1. to show that $\min(A, f)$ is not strongly well-posed. Let $\{x_n, y_n\} = \{n, 1/n\}_{n\in\mathbb{N}}$ be a sequence in X. We can see that

$$f(x_n, y_n) = -1, \quad \forall n \in \mathbb{N},$$

thus $\limsup_{n\to\infty} f(x_n, y_n) \leq \inf_A f$ and

$$(n, \tfrac{1}{n}) \to (\infty, 0) \in A \Rightarrow d((x_n, y_n), A) \to 0, \quad n \to \infty,$$

therefore $\{n, 1/n\}$ is a strongly minimizing sequence, but it does not converge to our solution $(0, 0)$, therefore $\min(A, f)$ is not s.w.p.

The notion of well-posedness in metric spaces is quite general and often requires stronger assumptions for the function f. Hence the setting of normed vector spaces, particularly Banach spaces combined with convexity, has been extensively researched. Including contributions by Lucchetti [2], Dontchev, Zolezzi [3] and many others. Working in such setting provides significant advantages, as proving existence and uniqueness of the solution is often more straightforward, because convexity simplifies working with minimizing sequences and the structure of Banach spaces offers the application of more advanced theorems. We will present an interesting result from Lucchetti [2], where strong well-posedness is assumed and under certain conditions we can explicitly approximate the optimal solution.

Definition 1.4.3. Let X be a vector space equipped with a norm $||\cdot|| : X \to \mathbb{R}$, then X is called a *Banach space* if it is complete.

Definition 1.4.4. Let X be a vector space, then a subset $C \subset X$ is said to be *convex* if

$$\forall x, y \in C : \lambda x + (1 - \lambda)y \in C, \quad \forall \lambda \in (0, 1).$$

Then the function $f : C \to \overline{\mathbb{R}}$ is called *convex* if

$$\forall x, y \in C, x \neq y : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall \lambda \in (0, 1).$$

Theorem 1.4.1. *Let X be a Banach space and $f \in \mathcal{F}(X)$ be l.s.c. such that $\lim_{||x||\to\infty} f = \infty$. Further let $g \in \mathcal{F}(X)$ be a convex function and suppose there exists $x \in X$ such that $g(x) < 0$, then put $A := \{x \in X : g(x) \leq 0\}$. Assume that the minimization problem* $\min(A, f)$ *is s.w.p. and $f_n(x)$ is constructed as such*

$$f_n(x) := f(x) + n \max\{g(x), 0\}.$$

Let $\{x_n\}$ be a sequence in X and $\{\varepsilon_n\}$ in $\mathbb{R}$ such that

$$f_n(x_n) \leq \inf_{x\in X} f_n(x) + \varepsilon_n, \quad \forall n \in \mathbb{N},$$

where $\{\varepsilon_n\}$ is a sequence in $\mathbb{R}$ such that $\varepsilon_n \downarrow 0$. Then $x_n \to x_0$, where $x_0 \in A$ is the unique solution for $\min(A, f)$.

Proof. Refer to Proposition 10.1.16, Lucchetti [2]

□

1.5 Hadamard well-posedness

Lastly, we introduce the notion of well-posedness that laid the foundation for all others. Jacques Hadamard developed the classical idea of well-posedness while studying the behavior of differential equations. His definition consists of the commonly aforementioned requirements, the existence and uniqueness of the optimal solution and its continous dependance on the data. The main distinction, and the reason Hadamard well-posedness is more broadly applicable than other types, is that Hadamard's concept extends to the entire problem $\min(A, f)$. Whereas before, we were using minimizing sequences as approximations for the solution, here we consider the approximations (A_n, f_n) of the pair (A, f), which determines $\min(A, f)$. We require that the sequence of solutions to the approximate problems converges to the optimal solution for the original problem.

Definition 1.5.1. Let $A \in \mathcal{H}(X)$ and $f \in \mathcal{F}(X)$. A minimization problem $\min(A, f) \in \mathcal{H}(X) \times \mathcal{F}(X)$ is called Hadamard well-posed (briefly H.w.p.) if all following conditions hold:

(i) there exists a unique solution $x_0 \in A$, i.e., argmin (A, f) is nonempty and a singleton;

(ii) for $A_n \in \mathcal{H}(X), f_n \in \mathcal{F}(X), \quad \forall n \in \mathbb{N}$
every sequence of pairs $\{(A_n, f_n)\} \to (A, f), \quad n \to \infty$,
with respect to the product metric ν on $\mathcal{H}(X) \times \mathcal{F}(X)$, where

$$\nu((A_n, f_n), (A, f)) = [\rho_H^2(A_n, A) + e^2(f_n, f)]^{1/2};$$

(iii) every sequence $\{x_n\}$ such that $\forall n \in \mathbb{N} : x_n \in \text{argmin}(A_n, f_n)$, converges to x_0.

We again provide a Furi-Vignoli-style characterization, now aligned with Hadamard well-posedness.

Theorem 1.5.1. *Let X be a complete and bounded metric space, $A \in \mathcal{H}(X)$ and $f \in \mathcal{F}(X)$ be a continous function. Then for the minimization problem* $\min(A, f)$ *the following assertions are equivalent:*

(i) $\min(A, f)$ *is Hadamard well-posed;*

(ii) $\inf\limits_{a>0} \text{diam } L^a = 0$,

where L^a is the generalized level set for A.

Proof. Refer to Theorem 17, p.96, Dontchev, Zolezzi [3]

□

Assuming the prerequisites of Theorem 1.5.1., we can see that (ii) implies Tikhonov well-posedness of $\min(A, f)$ according to the Furi-Vignoli criterion. As a consequence of the relation between the level sets $f^{a+\inf_A f} = L^a$, when we are restricted only to A, as we are for Tikhonov well-posedness. In the following example from Example 19 [3], we will show that T.w.p. does not imply H.w.p., if f is only lower semicontinous.

Example 1.5.1. Let $X = A = [0, 1]$ and $f : X \to \mathbb{R}$ be defined as

$$f(x) = \begin{cases} 2 - x, & 0 < x \leq 1, \\ 0, & x = 0. \end{cases} \tag{1.2}$$

Clearly $\inf_A f = 0$ is a unique minimum point attained only for $x_0 = 0$. Also the closed interval $[0, 1]$ is a complete space and f is l.s.c., because $2 - x$ is continous and for every sequence $\{x_n\}$ in A such that $x_n \to 0$ we get

$$\liminf_{n\to\infty} f(x_n) \geq 0.$$

Let $a = \inf_A f + \varepsilon$ for $\varepsilon \in (0,1)$, then the level set $f^a = \{0\}$ and so $\text{diam} f^a = 0$, therefore $\min(A, f)$ is T.w.p. by applying the Furi-Vignoli criterion.
Now let $A_n = [\frac{1}{n}, 1]$, then $(A_n, f) \to (A, f)$ and $\text{argmin}(A_n, f) = \{1\}, \forall n \in \mathbb{N}$, because

$$f(x) = 2 - x, \quad \tfrac{1}{n} \leq x \leq 1$$

obtains $\inf_{A_n} f = 1$ for $x = 1$. Let $\{x_n\}$ be a sequence such that $x_n \in \text{argmin}(A_n, f)$ for all $n \in \mathbb{N}$, then $\{x_n\} = (1, \ldots, 1)$ and we see that x_n does not converge to the optimal solution $x_0 = 0$, hence $\min(A, f)$ is not H.w.p. while being T.w.p. for f l.s.c.

In the beginning of the 20th century, when Hadamard introduced the notion of well-posedness, he presented the Cauchy problem for Laplace equation as an example of an ill-posed problem, and as to why it is important to research the concept of well-posedness. We will show the version of the historically important example as in Example 3.12, Kabanikhin [6].

Example 1.5.2. Cauchy problem consists of finding a solution of a partial differential equation that satisfies certain conditions, specifically in our case the Laplace equation, which is a second-order partial differential equation for $u = u(x_1, \ldots, x_n)$ formulated as

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = 0. \tag{1.3}$$

Our specific example of the Cauchy problem for Laplace equation is defined in the following way. Let $x > 0$ and $y \in \mathbb{R}$, then the problem consists of finding $u = u(x, y)$ for given data $f(y)$ such that all following conditions are satisfied

$$\Delta u = 0, \quad u(0, y) = f(y), \quad \frac{\partial u}{\partial x}(0, y) = 0. \tag{1.4}$$

Let the data $f(y)$ be chosen for $n \in \mathbb{N}$ in the following way

$$f(y) = u(0, y) = \tfrac{1}{n} \sin(ny).$$

After substitution into the Laplace equation and solving the ordinary differential equations we get the general solution for $c_1, c_2, d_1, d_2 \in \mathbb{R}$

$$u(x, y) = \left(c_1 e^{nx} + c_2 e^{-nx}\right) \left(d_1 \cos(ny) + d_2 \sin(ny)\right).$$

By applying the conditions (1.4) we obtain the solution to the problem

$$u(x, y) = \tfrac{1}{n} \sin(ny) \left(e^{nx} + e^{-nx}\right), \quad \forall n \in \mathbb{N}.$$

The essential concept of Hadamard well-posedness is that the solution must continously depend on the data, which means that small perturbations in the data should not change the solution by a large margin. In our example we see that for any fixed $x > 0$, the solution $u(x, y)$ reaches large values as $n \to \infty$, while the data $f(y)$ approaches 0. Therefore, small changes in the data can lead to indefinitely large changes in the solution. As a result the problem is ill-posed in Hadamard sense.

1.6 Relations of well-posedness

In this section, we determine the relations among the four types of well-posedness introduced earlier with the corresponding proofs. Before proceeding, we first define the concept of upper semicontinuity for functions.

Definition 1.6.1. A function $f \in \mathcal{F}(X)$ is said to be *upper semicontinous* (briefly u.s.c.) at $x_0 \in X$ if $\forall \varepsilon > 0$, there exists $\delta > 0$ such that

$$f(x_0) + \varepsilon > f(x) : \forall x \text{ that satisfy } d(x, x_0) < \delta$$

f is called *upper semicontinous* if f is u.s.c. at every point of its domain.
In our context the following equivalence might be more telling

$$f \text{ is u.s.c. at } x_0 \Leftrightarrow \forall \{x_n\} \in X, x_n \to x_0 : \limsup_{n\to\infty} f(x_n) \leq f(x_0).$$

We also need to establish that $\inf(\cdot,\cdot) : \mathcal{H}(X) \times \mathcal{F}(X) \to \mathbb{R}$ is u.s.c., provided f is u.s.c. It is proved in [4],Lemma 1.3, for the case that f is continous, so we need to prove for f u.s.c. In similar fashion we also need to prove that $\inf(\cdot,\cdot)$ is continous, if f is uniformly continous. The ensuing lemmas will be important in the proof of Theorem 1.6.1.

Lemma 1.6.1. *Let $\mathcal{F}_u(X)$ be a set of all functions $f \in \mathcal{F}(X)$ that are upper semicontinous. Let $\inf(\cdot,\cdot) : \mathcal{H}(X) \times \mathcal{F}_u(X) \to \mathbb{R}$ be a function, where $\inf(A, f) = \inf_A f$, then it is upper semicontinous everywhere in $\mathcal{H}(X) \times \mathcal{F}_u(X)$.*

Proof. Given $(A_0, f_0) \in \mathcal{H}(X) \times \mathcal{F}_u(X)$ and $\varepsilon > 0$, there exists $x_0 \in A_0$ such that

$$f_0(x_0) \leq \inf_{A_0} f_0 + \varepsilon/3.$$

We can afford to use $\leq$ instead of $<$ in most cases, because all $A \in \mathcal{H}(X)$ are closed subsets of X. Upper semicontinuity of f_0 provides the existence of $\delta \in (0, 1)$ such that $\frac{\delta}{1-\delta} \leq \varepsilon/3$, and

$$f_0(x) - f_0(x_0) \leq \varepsilon/3, \quad \forall x \in \{x \in A_0 : d(x, x_0) \leq \delta\}.$$

For $f \in B(f_0, \delta)$, where $B(f_0, \delta) = \{f \in \mathcal{F}_u(X) : e(f, f_0) < \delta\}$, we have

$$e(f, f_0) = \sup_{x\in X} \left\{ \frac{|f(x)-f_0(x)|}{1+|f(x)-f_0(x)|} \right\} < \delta,$$

which means that

$$f(x) - f_0(x) < \frac{\delta}{1-\delta} \leq \varepsilon/3, \quad x \in X.$$

Let $(A, f) \in B(A_0, \delta) \times B(f_0, \delta)$, where $B(A_0, \delta) = \{A \in \mathcal{H}(X) : \rho_H(A, A_0) < \delta\}$. Since $\rho_H(A, A_0) < \delta$ then there exists $x_1 \in A$ such that $d(x_1, x_0) \leq \delta$. Now we use all the aforementioned inequalities

$$\inf(A, f) = \inf_A f \leq f(x_1) \leq f_0(x_1) + \varepsilon/3 \leq f_0(x_0) + \varepsilon/3 + \varepsilon/3 \leq$$
$$\leq \inf_{A_0} f_0 + \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \inf_{A_0} f_0 + \varepsilon$$

which gets us $\inf(A_0, f_0)+\varepsilon \geq \inf(A, f), \quad \forall (A, f) \in B(A_0, \delta)\times B(f_0, \delta)$. Therefore $\inf(\cdot,\cdot)$ is u.s.c. at every point in $\mathcal{H}(X) \times \mathcal{F}_u(X)$.

□

Lemma 1.6.2. *Let $\mathcal{F}_{uc}(X)$ be a set of all functions $f \in \mathcal{F}(X)$ that are uniformly continous. Let $\inf(\cdot,\cdot) : \mathcal{H}(X) \times \mathcal{F}_{uc}(X) \to \mathbb{R}$ be a function, where $\inf(A, f) = \inf_A f$, then it is continous everywhere in $\mathcal{H}(X) \times \mathcal{F}_{uc}(X)$.*

Proof. Given $(A_0, f_0) \in \mathcal{H}(X) \times \mathcal{F}_{uc}(X)$ we will first establish the implications of f being uniformly continous.
Let $f \in \mathcal{F}_{uc}(X)$ and $\varepsilon > 0$, then there exists $\delta > 0$ such that for $\forall A \in B(A_0, \delta)$, where $B(A_0, \delta) = \{A \in \mathcal{H}(X) : \rho_H(A, A_0) < \delta\}$, we get the following according to the uniform semicontinuity of f

$$\forall x \in A, \forall x_0 \in A_0 : d(x, x_0) < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon/2,$$

and consequently

$$|\inf(A, f) - \inf(A_0, f)| < \varepsilon/2.$$

Now choose $\delta_\varepsilon \in (0, 1)$ such that $\frac{\delta_\varepsilon}{1-\delta_\varepsilon} < \varepsilon/2$. For $f \in B(f_0, \delta_\varepsilon)$, where $B(f_0, \delta_\varepsilon) = \{f \in \mathcal{F}_{uc}(X) : e(f, f_0) < \delta_\varepsilon\}$, we have

$$e(f, f_0) = \sup_{x \in X} \left\{ \frac{|f(x)-f_0(x)|}{1+|f(x)-f_0(x)|} \right\} < \delta_\varepsilon$$

which means that

$$f(x) - f_0(x) < \frac{\delta_\varepsilon}{1-\delta_\varepsilon} < \varepsilon/2, \quad \forall x \in A_0,$$

then

$$|\inf(A_0, f) - \inf(A_0, f_0)| < \varepsilon/2.$$

Put together we get

$$|\inf(A, f) - \inf(A_0, f_0)| \le |\inf(A, f) - \inf(A_0, f)| + |\inf(A_0, f) - \inf(A_0, f_0)| <$$

$$< \varepsilon/2 + \varepsilon/2 < \varepsilon.$$

Hence $|\inf(A, f) - \inf(A_0, f_0)| < \varepsilon$ for $\forall (A, f) \in B(A_0, \delta_\varepsilon) \times B(f_0, \delta_\varepsilon)$. Therefore $\inf(\cdot,\cdot)$ is continous at every point in $\mathcal{H}(X) \times \mathcal{F}_{uc}(X)$.

□

At last, we present the theorem and its expanded proof regarding the relations between types of well-posedness, as described in Theorem 2.1,Revalski,Zhivkov [5].

Theorem 1.6.1. *, Let $(A, f) \in \mathcal{H}(X) \times \mathcal{F}(X)$. Consider the following assertions:*

(i) $\min(A, f)$ *is Hadamard well-posed;*

(ii) $\min(A, f)$ *is strongly well-posed;*

(iii) $\min(A, f)$ *is Levitin-Polyak well-posed;*

(iv) $\min(A, f)$ *is Tikhonov well-posed.*

Then the ensuing implications are true, (i) ⇒ (ii) ⇒ (iii) ⇒ (iv).
Assuming additional conditions, we obtain the following equalities.
If f is upper semicontinous, then (i) ⇔ (ii).
If X is a normed vector space and f is continous, then (ii) ⇔ (iii).
If f is uniformly continous, then (i) ⇔ (iv).

Proof.
(i) ⇒ (ii); Assume that $\min(A, f)$ is H.w.p. with a unique solution $x_0 \in A$. Let $\{x_n\}$ be a strongly minimizing sequence of $\min(A, f)$. We will now construct a sequence of pairs $\{(A_n, f_n)\}$ and show that it complies with the definition of Hadamard well-posedness. Let $M \subseteq \mathbb{N}$ be a subset of natural numbers denoted as $M := \{n \in \mathbb{N} : f(x_n) \leq \inf_A f\}$ and a complement set $M^C = \mathbb{N} \backslash M$. Then we construct $f_n(x)$ for $x \in X$ in the following manner:

$$\forall n \in M : f_n(x) = f(x),$$

$$\forall n \in M^C : f_n(x) = \begin{cases} f(x) - \varepsilon_n, & f(x) - f(x_n) \geq \varepsilon_n, \\ f(x_n), & |f(x) - f(x_n)| < \varepsilon_n, \\ f(x) + \varepsilon_n, & f(x) - f(x_n) \leq -\varepsilon_n, \end{cases} \tag{1.5}$$

where $\varepsilon_n = f(x_n) - \inf_A f$.
We can see that for $\forall n \in M^C : f(x_n) > \inf_A f$, therefore $\varepsilon_n > 0$. This construction of f_n ensures that f_n gets closer to f as $f(x_n)$ approaches $\inf_A f$. We get $e(f_n, f) = 0$ for $n \in M$ and $e(f_n, f) \leq \varepsilon_n$ for $n \in M^C$. Applying the s. minimizing property of $\{x_n\}$, $\limsup f(x_n) \leq \inf_A f$, we get that f_n converges to f with respect to the e metric in $\mathcal{F}(X)$ and since f is bounded from below, accordingly f_n is also bounded from below and so $f_n \in \mathcal{F}(X)$ for all $n \in \mathbb{N}$. Let $A_n \in X$ be defined as $A_n = A \cup \{x_n\}$ for all $n \in \mathbb{N}$. Then $x_n \in \operatorname{argmin}(A_n, f_n)$, because $f_n(x) \geq f_n(x_n), \forall x \in A_n$. From the minimizing property of $\{x_n\}$ we use $d(x_n, A) \to 0$ and get that $A_n \to A$ in the Hausdorff metric ρ_H. Therefore $(A_n, f_n) \to (A, f)$ and we get H.w.p of $\min(A, f)$ with the unique solution x_0, and thus $x_n \to x_0$, hence $\min(A, f)$ is s.w.p.

(ii) ⇒ (iii); For this proof we can compare strongly and Levitin-Polyak minimizing sequences. Assume a sequence $\{x_n\}$ in X and $x_n \to x_0$, then for said sequence $\limsup f(x_n) \leq \inf_A f$ is a weaker condition than $f(x_n) \to \inf_A f$, where $\inf_A f = f(x_0)$. So if we define $P_s \subset \mathcal{H}(X) \times \mathcal{F}(X)$ as a set of all s.w.p. minimization problems $\min(A, f)$ and likewise P_{LP} for all L.P.w.p. problems, then $P_{LP} \subset P_s$ and as a result s.w.p implies L.P.w.p.

(iii) ⇒ (iv); We can view T.w.p. of $\min(A, f)$ as a generalization of L.P.w.p. considering the behavior of f only for $x \in A$. Assume that $\min(A, f)$ is L.P.w.p., then every L.P. minimizing sequence converges to a unique solution $x_0 \in A$. Let $\{x_n\} \in A$ be L.P. minimizing for $\min(A, f)$, then for all $n \in \mathbb{N}$, we get $\mathrm{d}(x_n, A) = 0$ and $f(x_n) \to \inf_A f$ which means that $\{x_n\}$ is also T. minimizing for $\min(A, f)$ and converges to x_0, therefore $\min(A, f)$ is T.w.p. and (iii) ⇒ (iv).

(ii) ⇒ (i); Suppose $\min(A, f)$ is s.w.p. with a unique solution $x_0 \in A$. Let $\{x_n\}$ be a sequence in X, $A_n \in \mathcal{H}(X)$ and $f_n \in \mathcal{F}(X)$ such that $(A_n, f_n) \to (A, f)$ and $x_n \in \operatorname{argmin}(A_n, f_n)$ for all $n \in \mathbb{N}$. Let f be u.s.c. and $\varepsilon > 0$, then there exists sufficiently large $n_0 \in \mathbb{N}$ such that

$$f_n(x_n) = \inf_{A_n} f_n \leq \inf_A f + \varepsilon/2, \quad \forall n \geq n_0,$$

which we get from $\inf(\cdot, \cdot)$ being u.s.c. according to Lemma 1.6.1. by the alternate definition of u.s.c, which yields that for all (A_n, f_n) such that

$$(A_n, f_n) \to (A, f), \text{ then } \limsup_{n\to\infty} \inf(A_n, f_n) \leq \inf(A, f).$$

Since $(A_n, f_n) \to (A, f)$ means $f_n \to f$ with respect to the e metric in $\mathcal{F}(X)$, consequently we get

$$f(x_n) \leq f_n(x_n) + \varepsilon/2 \leq \inf_A f + \varepsilon, \quad \forall n \geq n_0.$$

From $(A_n, f_n) \to (A, f)$ we get that $A_n \to A$ in ρ_H metric, therefore $d(x_n, A) \to 0$. As a result $\{x_n\}$ is a strongly minimizing sequence for $\min(A, f)$, which gives us $x_n \to x_0$ and that means $\min(A, f)$ is also H.w.p. with the unique solution x_0.

(iii) $\Rightarrow$ **(ii)**; Let $\min(A, f)$ be L.P.w.p with a unique solution $x_0 \in A$. Assume X is a normed vector space and f is continous. Let $\{x_n\}$ be a strongly minimizing sequence for $\min(A, f)$, then we get

$$d(x_n, A) \to 0 \text{ and } \limsup f(x_n) \leq \inf_A f,$$

where $d(x_n, A) = \inf\{||x_n - y|| : y \in A\}$, because X is a normed vector space. Since $d(x_n, A) \to 0$, then there exists a sequence $\{y_n\}$ in A such that $||x_n - y_n|| \to 0$. For all $n \in \mathbb{N}$ we have $f(y_n) \geq \inf_A f$ and without loss of generality assume that $f(x_n) \to \lambda$ and $f(y_n) \to \mu$, then we get

$$\lambda \leq \inf_A f \leq \mu.$$

Since f is bounded from below and continous, hence the interval $[\lambda, \mu]$ contains values of f. Now according to the Intermediate value theorem for continous functions, there exists a sequence $\{z_n\}$ in X for any $\varepsilon > 0$ such that $||z_n - y_n|| \leq ||x_n - y_n||$ and $|f(z_n) - \inf_A f| < \varepsilon$ as $n \to \infty$, hence

$$f(z_n) \to \inf_A f, \quad \varepsilon \to 0.$$

Therefore $\{z_n\}$ is L.P. minimizing, which means $z_n \to x_0$, because $\min(A, f)$ is L.P.w.p. as assumed. Now we show that $z_n \to x_0$ implies $x_n \to x_0$,

$$||x_n - z_n|| \leq ||x_n - y_n|| + ||y_n - z_n|| \leq 2||x_n - y_n|| \to 0.$$

Consequently every strong minimizing sequence $\{x_n\}$ converges to the unique solution $x_0 \in A$, thereby $\min(A, f)$ is s.w.p.

(iv) $\Rightarrow$ **(i)**; Assume that a minimization problem $\min(A, f)$ is T.w.p. with a unique solution $x_0 \in A$ and f is uniformly continous. Let $\{x_n\}$ be a sequence in X, $A_n \in \mathcal{H}(X)$ and $f_n \in \mathcal{F}(X)$ such that $(A_n, f_n) \to (A, f)$, where $x_n \in \operatorname{argmin}(A_n, f_n)$ for all $n \in \mathbb{N}$. Then according to Lemma 1.6.2 the $\inf(\cdot, \cdot)$ function is continous for f uniformly continous, which gives

$$(A_n, f_n) \to (A, f) \Rightarrow \inf_{A_n} f_n \to \inf_A f.$$

Since $x_n \in \operatorname{argmin}(A_n, f_n)$, then $f_n(x_n) = \inf_{A_n} f_n$ and $f(x_0) = \inf_A f$, because $\min(A, f)$ is T.w.p. with the unique solution x_0. Now using that $f_n \to f$ we get

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f_n(x_n) = \lim_{n\to\infty} \inf_{A_n} f_n = \inf_A f = f(x_0),$$

therefore $f(x_n) \to f(x_0)$, and thus $\{x_n\}$ is T. minimizing sequence for $\min(A, f)$, which yields $x_n \to x_0$. As a result $\min(A, f)$ is H.w.p.

This concludes the proof of all implications in Theorem 1.6.1. □

2 General well-posedness

2.1 Characterization of general well-posedness

In the previous chapter, we introduced three key conditions for well-posedness and the uniqueness of the optimal solution was one of them. Yet, in some contexts, this assumption can be too restrictive or even unnecessary. Therefore, we now relax this requirement and extend the concept of well-posedness to minimization problems that may not have a unique solution, introducing the notion of general well-posedness. Recall that the previously defined types of well-posedness all required every minimizing sequence of approximate solutions to converge to the optimal solution. With the uniqueness requirement omitted, it is now sufficient for each minimizing sequence to have a subsequence that converges to one of the optimal solutions. Here, we will define the general versions of the four previously introduced types of well-posedness, incorporating these relaxed requirements.

Consider X a metric space and $f \in \mathcal{F}(X)$ unless mentioned otherwise. We will start with the definition of the generalized Tikhonov well-posedness for an unconstrained minimization problem $\min(X, f)$.

Definition 2.1.1. A minimization problem $\min(X, f)$ is *generalized Tikhonov well-posed* (briefly g.T.w.p.) if all following conditions hold:

(i) $\operatorname{argmin}(X, f) \neq \emptyset$;

(ii) every minimizing sequence of $\min(X, f)$ has a subsequence converging to $x_0 \in \operatorname{argmin}(X, f)$.

One of the changes in approach is that, before, we focused on whether $\operatorname{argmin}(X, f)$ was a singleton for the problem $\min(X, f)$. In our new setting, we will usually be interested in the compactness of $\operatorname{argmin}(X, f)$. From the Definition 2.0.1., we can see that if $\min(X, f)$ is g.T.w.p., then $\operatorname{argmin}(X, f)$ is compact, since it contains the limits of sequences in X.

Definition 2.1.2. A subset $A \subset X$ is *compact* if for every collection C of open subsets of X such that

$$A \subseteq \bigcup_{S \in C} S,$$

there exists a finite subcollection $D \subseteq C$ such that

$$A \subseteq \bigcup_{S \in D} S.$$

Another formulation holds in metric spaces, which might fit our context better. The subset $A \subset X$ is compact if every infinite subset of A has a limit point in A.

Next is a simple example showing the difference between T.w.p. and g.T.w.p. inspired by Examples 34, Dontchev, Zolezzi [3].

Example 2.1.1. Let $X = \mathbb{R}$ and $f(x) = |x| - a$ for $a \in \mathbb{R}^+$, then we can see that $\inf_X f = -a$ attained only for $x = 0$, therefore we have a unique solution $x_0 = 0$ for the minimization problem $\min(X, f)$. Let $\{x_n\}$ be any T. minimizing sequence, then $|x_n| - a \to -a$ and hence $x_n \to 0 = x_0$. We get that $\min(X, f)$ is T.w.p.

Now consider the modification to the problem, let $f(x) = ||x| - a|$. In this case $\inf_X f = 0$ for $x = a$ and $x = -a$, therefore $\operatorname{argmin}(X, f) = \{-a, a\}$, which is not a singleton, hence $\min(X, f)$ cannot be T.w.p. But let $\{x_n\}$ be any T. minimizing sequence, then $||x_n| - a| \to 0$ and so $|x_n| \to a$. We end up with two cases, either $x_n \to a$ or $x_n \to -a$. Without loss of generality consider the first case $x_n \to a$, then we can take a subsequence $\{x_k\}$, where $x_k > 0$, $\forall k$ and we get that $x_k \to a$ since $a \in \mathbb{R}^+$. As a result the modified problem $\min(X, f)$ is generalized T.w.p. while not being T.w.p.

We now introduce the concept of Furi-Vignoli criterion for generalized Tikhonov well-posedness.

Theorem 2.1.1. *Let X be a complete metric space and $f \in \mathcal{F}(X)$ be lower semicontinous, then the following assertions are equivalent:*

(i) $\min(X, f)$ *is Tikhonov well-posed in the generalized sense;*

(ii) $\operatorname{argmin}(X, f)$ *is compact and*

$$\forall \varepsilon > 0,\ \exists\, a > \inf_X f\ :\ f^a \subset B\left[\operatorname{argmin}(X, f), \varepsilon\right],$$

where $B\left[\operatorname{argmin}(X, f), \varepsilon\right] = \{A \in X : \rho_H(A, \operatorname{argmin}(X, f)) \leq \varepsilon\}$.

Proof. Refer to Proposition 10.1.7, Lucchetti [2] □

For constrained problems $\min(A, f)$, where $A \subset X$, we get that the generalized versions of Levitin-Polyak and strong well-posedness follow the same principle as g.T.w.p. in regard to the relaxation of the uniqueness requirement.

Definition 2.1.3. A minimization problem $\min(A, f)$ is *generalized Levitin-Polyak well-posed* (briefly g.L.P.w.p.) if all following conditions hold:

(i) $\operatorname{argmin}(A, f) \neq \emptyset$;

(ii) every Levitin-Polyak minimizing sequence of $\min(A, f)$ has a subsequence converging to $x_0 \in \operatorname{argmin}(A, f)$.

Definition 2.1.4. A minimization problem $\min(A, f)$ is *generalized strongly well-posed* (briefly g.s.w.p.) if all following conditions hold:

(i) $\operatorname{argmin}(A, f) \neq \emptyset$;

(ii) every strongly minimizing sequence of $\min(A, f)$ has a subsequence converging to $x_0 \in \operatorname{argmin}(A, f)$.

Eeach type of mentioned well-posedness can charactarized according to the concept of the Furi-Vignoli criterion, but in the case of general well-posedness, it is slightly more complex. Since the characterization of s.w.p. in the previous chapter was omitted, we will now mention it here for g.s.w.p. as in Theorem 4.1, Revalski, Zhivkov [5]. For that we will need to define the Kuratowski measure of noncompactness.

Definition 2.1.5. Let X be a metric space and C_d a collection of subsets in X such that for every $S \in C_d$: diam $S \leq d$, then the *Kuratowski measure of noncompactness* for $A \subset X$ is defined as

$$\alpha(A) = \inf \left\{ d > 0 : C_d \text{ is finite and } A \subseteq \bigcup_{S \in C_d} S \right\}.$$

Theorem 2.1.2. *Let $\emptyset \neq A \subset X$, $f \in \mathcal{F}(X)$ and L^a be a generalized level set for A, then the following assertions hold:*

(i) If $\min(A, f)$ *is g.s.w.p., then* $\alpha(L^a) \to 0$ *as* $a \to 0$;

(ii) If X is a complete, $A \in \mathcal{H}(X)$, f is l.s.c. and $\alpha(L^a) \to 0$ as $a \to 0$, then $\min(A, f)$ *is g.s.w.p.*

Proof. Refer to Theorem 4.1, Revalski, Zhivkov [5] □

In Theorem 4.3, Revalski, Zhivkov [5] it is shown that we can obtain equivalence between g.L.P.w.p. and g.s.w.p. Let $f \in \mathcal{F}(X)$ be continous and for every two sequences $\{x_n\}, \{y_n\}$ in X such that $d(x_n, y_n) \to 0$ and $f(x_n) \to \lambda$, $f(y_n) \to \mu$, where $\lambda, \mu \in \mathbb{R}$. Additionaly for every $p \in (\lambda, \mu)$, there exists a sequence $\{z_n\}$ in X such that $d(x_n, z_n) \to 0$ and p is a cluster point of $\{f(z_n)\}$, which means that for every $\varepsilon > 0$ there exists an infinite subsequence $\{z_{n_k}\}$ such that $|f(z_{n_k}) - p| < \varepsilon$. If the aforementioned properties are satisfied, then for every nonempty $A \subset X$ the minimization problem $\min(A, f)$ is g.L.P.w.p. if and only if it is g.s.w.p.

Lastly, from Revalski, Zhivkov [5], we present the notion of generalized Hadamard well-posedness and the generalized version of Theorem 1.6.1.

Definition 2.1.6. Let $A \in \mathcal{H}(X)$ and $f \in \mathcal{F}(X)$. We define the set of solutions argmin as a multivalued mapping argmin : $\mathcal{H}(X) \times \mathcal{F}(X) \to X$. A minimization problem $\min(A, f)$ is *generalized Hadamard well-posed* (briefly g.H.w.p.) if all following conditions hold:

(i) argmin is upper semicontinous;

(ii) $\operatorname{argmin}(A, f) \neq \emptyset$ and is a compact subset of X.

Theorem 2.1.3. *Let $(A, f) \in \mathcal{H}(X) \times \mathcal{F}(X)$. Consider the following assertions:*

(i) $\min(A, f)$ *is g.H.w.p.;*

(ii) $\min(A, f)$ *is g.s.w.p.;*

(iii) $\min(A, f)$ *is g.L.P.w.p.;*

(iv) $\min(A, f)$ *is g.T.w.p.*

Then the ensuing implications are true, (i) ⇒ (ii) ⇒ (iii) ⇒ (iv).
Assuming additional conditions, we obtain the following equalities.
If f is upper semicontinous, then (i) ⇔ (ii).
If f is uniformly continous, then (i) ⇔ (iv).

Proof. Refer to Theorem 4.2, Revalski, Zhivkov [5]

□

3 Numeric example

3.1 Two-asset minimum-variance portfolio

Consider the problem of minimizing the variance of a two-asset portfolio. Let $X, Y \in \mathbb{R}^n$ be the performance data of the two assets respectively over some time period $n \in \mathbb{N}$. For example consider the yearly performance of two stock options over 10 year period. We seek to find an optimal solution of sharing our wealth between the two assets such that the volatility of our investment is minimal. As a main result, we will prove that said problem is well-posed in every sense.

To begin we denote the fundamental characteristics of the data.
Expected return of each asset:

$$R_X = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad R_Y = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

Sample variance of each asset:

$$\sigma_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2, \quad \sigma_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}_n\right)^2$$

Sample covariance of the two assets:

$$\sigma_{XY} = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)\left(Y_i - \overline{Y}_n\right)\right]$$

Let $\omega_X, \omega_Y \in [-1, 1]$ be the portfolio weights for each asset respectively. Positive values denote long positions and negative values denote short positions. We assume that all wealth is invested, therefore $\omega_X + \omega_Y = 1$. For given weights we can define the *expected portfolio return* as such:

$$R_p = \omega_X \cdot R_X + \omega_Y \cdot R_Y \tag{3.1}$$

The following equation denotes the variance of the portfolio:

$$\sigma_p^2 = \omega_X^2 \cdot \sigma_X^2 + \omega_Y^2 \cdot \sigma_Y^2 + 2 \cdot \omega_X \cdot \omega_Y \cdot \sigma_{XY}$$

Applying $\omega_X + \omega_Y = 1$, we get a more suitable formulation of the equation:

$$\sigma_p^2 = \omega_X^2(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}) + \omega_X(-2\sigma_Y^2 + 2\sigma_{XY}) + \sigma_Y^2 \tag{3.2}$$

Theorem 3.1.1. *Let $A = [-1, 1], \omega_X \in A$ and σ_p^2 be a function as in (3.6) determined by the data $X, Y \in \mathbb{R}^n$, then the two-asset portfolio variance minimization problem* $\min(A, \sigma_p^2)$,

$$\min_{\omega_X} \sigma_p^2, \tag{3.3}$$

is well-posed in every sense mentioned in Theorem 1.6.1.

Proof. First we need to show that a solution exists and is unique for

$$\sigma_p^2 = \omega_X^2(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}) + \omega_X(-2\sigma_Y^2 + 2\sigma_{XY}) + \sigma_Y^2.$$

We will search for a global minimum point in A using the first derivative of σ_p^2,

$$\frac{\partial \sigma_p^2}{\partial \omega_X} = 2\omega_X(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}) - 2\sigma_Y^2 + 2\sigma_{XY} = 0.$$

As a result we get the only extreme point,

$$\omega_X = \frac{\sigma_Y^2 - \sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}}. \tag{3.4}$$

Continuing with the second derivative to determine if ω_X is a minimum point of σ_p^2,

$$\frac{\partial^2 \sigma_p^2}{\partial \omega_X^2} = 2(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}),$$

$$\frac{\partial^2 \sigma_p^2}{\partial \omega_X^2} = \frac{2}{n-1}\left[\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2 - 2\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)\left(Y_i - \overline{Y}_n\right) + \sum_{i=1}^{n}\left(Y_i - \overline{Y}_n\right)^2\right],$$

$$\sum_{i=1}^{n}\Big((X_i + Y_i) - (\overline{X}_n + \overline{Y}_n)\Big)^2 > 0 \Rightarrow \frac{\partial^2 \sigma_p^2}{\partial \omega_X^2} > 0. \tag{3.5}$$

The second derivative is positive for all $\omega_X \in A$, hence the function σ_p^2 is convex in A and so ω_X is a global minimum point of σ_p^2. Furthermore the inequality in (3.5) gives us that the denominator in (3.4) is never zero, therefore the solution always exists and, as mentioned before, is unique. Since σ_p^2 attains its minimum value in A, then σ_p^2 is bounded from below and proper, hence $\sigma_p^2 \in \mathcal{F}(A)$. Let $\{\omega_{X_n}\}$ be a T. minimizing sequence in A, then through the series of the following operations,

$$\sigma_p^2(\omega_{X_n}) \to \sigma_p^2(\omega_X) \Leftrightarrow \sigma_p^2(\omega_{X_n}) - \sigma_p^2(\omega_X) \to 0,$$

$$(\omega_{X_n}^2 - \omega_X^2)(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}) + (\omega_{X_n} - \omega_X)(-2\sigma_Y^2 + 2\sigma_{XY}) \to 0,$$

$$\frac{\omega_{X_n}^2 - \omega_X^2}{\omega_{X_n} - \omega_X} \to \frac{2\sigma_Y^2 - 2\sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}} = 2\omega_X,$$

$$\omega_{X_n} + \omega_X \to 2\omega_X \Leftrightarrow \omega_{X_n} \to \omega_X,$$

we obtain $\omega_{X_n} \to \omega_X$, which yields that $\min(A, \sigma_p^2)$ is T.w.p.

To establish other types of well-posedness, we must show that the function σ_p^2 is uniformly continuous on A. Let $a, b \in [-1, 1]$, then $-2 \leq a + b \leq 2$ and we get

$$|\sigma_p^2(a) - \sigma_p^2(b)| \leq |a - b| \Big[|(a+b)(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY})| + | - 2\sigma_Y^2 + 2\sigma_{XY}| \Big] \leq$$

$$\leq \Big[|2(\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY})| + | - 2\sigma_Y^2 + 2\sigma_{XY}| \Big] \cdot |a - b|.$$

Therefore σ_p^2 is Lipschitz continous on A, which implies that σ_p^2 is uniformly continous on A. We can clearly see that A is a closed set, consequently we can use Theorem 1.6.1, specifically that T.w.p. implies H.w.p., which in turn implies both L.P.w.p. and s.w.p. Altogether we have proved that the two-asset portfolio variance minimization problem $\min(A, \sigma_p^2)$ is well-posed in every sense.

□

We will now examine the continous dependancy of the expected portfolio return function R_p on the data with a simple example. Let $\varepsilon_X, \varepsilon_Y \in \mathbb{R}$ be a some kind of fixed errors in the data X, Y such that $X_i^\varepsilon = X_i + \varepsilon_X$ and $Y_i^\varepsilon = Y_i + \varepsilon_Y$ for $i \in \{1, \ldots, n\}$, then

$$R_X^\varepsilon = \overline{X}_n^\varepsilon = \frac{1}{n} \sum_{i=1}^{n} (X_i + \varepsilon_X) = R_X + \varepsilon_X, \; R_Y^\varepsilon = \overline{Y}_n^\varepsilon = \frac{1}{n} \sum_{i=1}^{n} (Y_i + \varepsilon_Y) = R_Y + \varepsilon_Y$$

$$\sigma_{X^\varepsilon}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i^\varepsilon - \overline{X}_n^\varepsilon \right)^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i + \varepsilon_X - \overline{X}_n - \varepsilon_X \right)^2 = \sigma_X^2.$$

We can see that the errors in this specific case do not affect the variance of X and the same goes for the variance of Y and their covariance, hence the solution ω_X is also unaffected. Thus we switch our focus to R_p, where the modified version is

$$R_p^\varepsilon = \omega_X \cdot R_X^\varepsilon + \omega_Y \cdot R_Y^\varepsilon = \omega_X (R_X + \varepsilon_X) + \omega_Y (R_Y + \varepsilon_Y) = R_p + \omega_X \varepsilon_X + \omega_Y \varepsilon_Y.$$

Using the substitution $\omega_X + \omega_Y = 1$, we get

$$R_p^\varepsilon = R_p + \omega_X (\varepsilon_X - \varepsilon_Y) + \varepsilon_X. \tag{3.6}$$

We can see a linear dependancy of the solution on the data, thefore such problem for R_p would be considered well-posed.

To illustrate with a practical example, let us consider a two-asset portfolio comprising Apple Inc. (AAPL) and ČEZ Group (CEZ) stocks. Assume we are provided with the historical data for these two stocks over the past five years. The data includes annualized expected returns, volatility and covariance over the 5Y period:

$$R_A = 0.330, \; \sigma_A^2 = 0.0831, \; R_C = 0.173, \; \sigma_C^2 = 0.0762, \; \sigma_{AC} = 0.00185$$

Plugging this data into $\min(A, \sigma_p^2)$ results in the following portfolio:

$$\omega_A = 0.0643, \; \omega_C = 0.9357, \; \sigma_p^2 = 0.0051, \; R_p = 0.1831.$$

4 Ill-posed problems

4.1 Ill-posed problems

Mathematical problems that fail to satisfy the criteria for well-posedness are referred to as ill-posed problems. These problems arise in various areas of mathematics. For instance, we previously encountered the Cauchy problem for the Laplace equation in Example 1.5.2, where small changes in the input data led to significant discrepancies in the solutions. This sensitivity makes the problem ill-posed in the Hadamard sense. Ill-posedness frequently appears in systems of linear equations, integral equations, partial differential equations, and many other mathematical frameworks. In physics, the study of inverse problems is crucial because such problems are often ill-posed. An inverse problem involves reconstructing the input data or causes based on observed outcomes. Examples include image reconstruction, the backward heat equation, audio signal restoration, for more see Chapter 5, Sizikov [7] and Chapter 3, Kabanikhin [6].

To illustrate the concept of an ill-posed problem, we present a simple example inspired by p.4, Sizikov, [7]. Consider the task of finding the intersection point of two "similar" lines in $\mathbb{R}^2$. Imagine a practical scenario where we attempt to draw two lines as close to each other as possible without making them identical or parallel. To locate their intersection, we measure the endpoints of each line and plot them in $\mathbb{R}^2$. Additionaly assume there exists a coefficient error ε, which corresponds to the error in our measuring.

Example 4.1.1. Consider a system of linear equations for $x, y \in \mathbb{R}$ given by data parameter $a \in \mathbb{R}$ and coefficient error $\varepsilon \in \mathbb{R}$:

$$\begin{cases} (a+0.1)x + y = (a+0.1), \\ (a+\varepsilon)x + y = a. \end{cases} \tag{4.1}$$

We can solve using Cramer's rule for determinants of coefficient matrices:

$$\det A = \begin{vmatrix} a+0.1 & 1 \\ a+\varepsilon & 1 \end{vmatrix}, \det A_1 = \begin{vmatrix} a+0.1 & 1 \\ a & 1 \end{vmatrix}, \det A_2 = \begin{vmatrix} a+0.1 & a+0.1 \\ a+\varepsilon & a \end{vmatrix}$$

and we get the solution

$$x = \frac{\det A_1}{\det A} = \frac{1}{1-10\varepsilon}, \quad y = \frac{\det A_2}{\det A} = \frac{-(a+0.1)\varepsilon}{0.1-\varepsilon}.$$

We observe that the solution for the x-coordinate is independent of the parameter a. Let us now examine how the coefficient error affects the solution.

$$\text{if } |\varepsilon| \leq 0.001, \text{ then } 0.99 \leq x \leq 1.01,$$
$$\text{if } |\varepsilon| \leq 0.01, \text{ then } 0.909 \leq x \leq 1.11,$$
$$\text{if } |\varepsilon| \leq 0.1, \text{ then } 0.5 \leq x \leq \infty.$$

It is evident that the solution increases rapidly, even with only minor changes in the coefficient error. Therefore the problem (3.1) is ill-posed in the Hadamard sense. The key takeaway is that the ill-posedness of solving systems of linear equations is closely tied to the determinant of the coefficient matrix. The smaller the determinant, the more ill-posed the problem becomes.

4.2 Tikhonov regularization method

Various methods can be employed to approximate a solution of an ill-posed problem. Tikhonov regularization is one such method, which further develops the well-known Gauss least-squares method. Tikhonov regularization gives the best approximate solution for the problem introduced in the following definition.

Definition 4.2.1. Let X, F be metric spaces and $A : X \to F$ be a linear operator from X into F. Consider the problem to find the solution $x \in X$ for given A and $f \in F$ such that

$$Ax = f, \quad x \in X, \quad f \in F. \tag{4.2}$$

Let $\tilde{A}$ and $\tilde{f} \in F$ be known approximations of said problem and assume the error $\varepsilon \geq 0$ in setting A and likewise $\delta > 0$ for f such that

$$||\tilde{f} - f|| \leq \delta,$$

$$||\tilde{A} - A|| \leq \varepsilon.$$

Then we define the approximate solution $\tilde{x} \in X$ such that

$$\tilde{A}\tilde{x} = \tilde{f}, \quad \tilde{x} \in X, \quad \tilde{f} \in F. \tag{4.3}$$

In the case of Example 3.1.1. the error in measuring ε corresponds to the error ε in setting the elements of the coefficient matrix A and δ would be an error in setting the right-hand side vector. For the problems defined by the Definition 3.2.1. a new notion of well-posedness is given.

Definition 4.2.2. Let X, F be metric spaces and $A : X \to F$ be a linear operator from X into F. The problem (3.4) is called *conditionally well-posed* if the following conditions hold:

(i) the solution x exists and belongs to a subset $M \subset X$;

(ii) A is invertible on the set M;

(iii) the inverse operator A^{-1} is continous.

If any of the conditions are not satisfied, then the problem is called *essentialy ill-posed.*

The known method to minimize the discrepancy $||Ax - f||$ is the Gauss Least Squares Method, where we solve the following

$$\min_x ||Ax - f||^2.$$

However, this approach often results in highly unstable solutions. To address this, Tikhonov introduced a generalized version incorporating a regularization parameter α that acts a Langrange multiplier. This method uses the same discrepancy minimization principle as the Gauss least-squares method, incorporating a regularization parameter to find a pseudo-solution. It then utilizes the Moore-Penrose pseudo-inverse matrix to obtain the desired solution.

Definition 4.2.3. A square matrix $A \in C^{n\times n}$ is called *Hermitian* if and only if

$$A = A^*,$$

where $A^* = \overline{A^T}$.

Definition 4.2.4. The *Moore-Penrose pseudo-inverse matrix* of a matrix $A \in C^{m\times n}$ is a matrix $A^\dagger \in C^{m\times n}$ that satisfies the following conditions:

$$AA^\dagger A = A,\ A^\dagger AA^\dagger = A^\dagger,\ (AA^\dagger)^* = AA^\dagger,\ (A^\dagger A)^* = A^\dagger A.$$

Definition 4.2.5. Let X, F be metric spaces and $A : X \to F$ be a linear operator from X into F. Consider the *regularization parameter* $\alpha > 0$, then the following version of the problem is *Tikhonov regularized problem*:

$$\min_x ||Ax - f||^2 + \alpha||x||^2.$$

Which leads to the *Euler-Tikhonov equation*

$$(\alpha I + A^*A)x_\alpha = A^*f,$$

with the ensuing solution,

$$x_\alpha = (\alpha I + A^*A)^{-1}A^*f,$$

where A^* is the Hermitian matrix of A and I is the unit operator, i.e., $Ix = x$.

We can see that if $\alpha = 0$, then the method is equal to the Gauss least-squares method. If $\delta, \varepsilon \to 0$, then $\alpha \to 0$ and we get a solution using the pseudo-inverse matrix.

$$x_\alpha = \lim_{\alpha\to 0} (\alpha I + A^*A)^{-1}A^*f \equiv A^\dagger f.$$

For further analysis of the method, see Subchapter 4.7, Sizikov [7]. The properties of α and its effect on the regularization are explored in Definition 4.23, Kabanikhin [6]. Examples of Tikhonov regularization method can be found in p.160-176, Sizikov [7].

5 Applied well-posedness

5.1 Physics and Engineering

In the last segment, we will highlight the practical applications of well-posedness. Each paragraph draws information from one or two specific research papers in the respective field, complemented by general knowledge.

Heat transfer The heat equation problem is similar to the Cauchy problem for Laplace equation as in Example 1.5.2., as it also involves a partial differential equation with boundary conditions. This problem consists of determining the temperature of an object at a given time based on the initial heat distribution. It is proved that the heat problem is ill-posed in reversed time, but well-posedness of the problem in forward time is of great interest. The heat equation serves as a mathematical model for heat diffusion, which describes how heat propagates through a material over time. The practical applications range from physical modeling to climate science. For example, in engineering, the heat equation is used in optimizing thermal systems in electronics. Here, well-posedness guarantees that an optimal solution exists and small measurement errors in initial temperatures do not result in largely inaccurate predictions.

Acoustics Acoustic wave equations are partial differential equations, which determine how pressure disturbances from acoustic waves propagate through a material. In practice, these equations are often solved using least-squares finite element methods. Applications of acoustic wave equations include optimizing noise-canceling systems, improving hearing aids, designing underwater sonar systems, advancing ultrasound imaging, and even analyzing seismic activity. Well-posedness is crucial, because it guarantees stability in the results.

Electromagnetic fields Maxwell's equations describe the propagation of electromagnetic waves in a polarizable medium. Predicting how electromagnetic fields interact is important in numerous fields. This enables the development of wireless networks, circuit designs in electronic devices, MRI scans, etc. Well-posedness provides the existence of the solution and uniqueness additionaly makes it possible to choose different appropriate methods to solve the corresponding problem. Precision in these systems is essential, therefore the stability given by well-posedness is indispensable.

Fluid dynamics In continuum mechanics, the Navier–Stokes equations are partial differential equations derived from the law of conservation of mass, describing the motion of viscous fluid substances. These equations are crucial in the design of power stations, the study of blood flow, and the design of aircrafts, etc. Again the stability provided by well-posedness is really important, as these problems have a significant impact on the functioning of society. The existence of solutions for the three-dimensional Navier–Stokes equations remains unproven. This concept is so important in physics that it has been deemed one of the seven Millennium Prize problems.

5.2 Finance

Forecasting option prices One of the most important concepts in modern financial theory is the Black-Scholes model. This model is used to try forecasting option prices. The Black-Scholes equation is solved in backward time, where current value of a financial option, interest rate, and volatility are considered. We solve for the strike price of the underlying asset and time. Proposed mathematical models using the Black-Scholes equation in forward time were shown to be accurate for 1-2 trading days ahead. In practice, the Black-Scholes models are far too volatile for longer time periods. It is proved that Black-Scholes model in forward time is ill-posed, therefore regularization methods are used to approximate the solution, namely quasi-reversibility method.

5.3 Machine learning

Learning Problem Machine learning is a field in artificial intelligence that consists of training machines to perform specific tasks without previous explicit programming. Learning problem is defined by the problem goal, data processing algorithm, and appropriate loss function, which measures the quality of learned data. Key characteristics of successful learning problem include a suitable model, relevant metrics, and enough quality input data. Well-posedness ensures that the quality of input data is passed on to the solution data and we avoid inefficient learning process.

5.4 Medicine

We have mentioned some applications of the previously introduced concepts in medicine, including the use of electromagnetic waves in MRI scans, acoustic waves to improve hearing aids, and fluid dynamics to study blood flow. It makes sense that well-posedness plays a big role in medicine, since any inaccuracies may lead to fatal mistakes.

Epidemiology Mathematical models in epidemiology are used to predict the spread of transmissive diseases. The model uses the non linear transport equation to simulate the spread. Determining parameters for this problem are incubation period and the transmission rate for susceptible individuals. The movement of infected individuals is simulated by space diffusion. Well-posedness provides reliable predictions for the spreading rate of the disease.

Conclusion

This thesis unified different notions of well-posed optimization problems from various sources. Our focus was on minimization problems in metric spaces. We introduced the concept of minimizing sequences and characterized each notion of well-posedness accordingly. The concept of Furi-Vignoli criterion was applied to most of the notions, therefore we have shown that level sets and minimization problems are closely related. Proving Lemma 1.6.1. and Lemma 1.6.2. provided additional properties for the $\inf(\cdot, \cdot)$ function, which led to the main result of this thesis. By completely proving Theorem 1.6.1., we have shown that the notions of well-posedness are equivalent in certain metric spaces. We hinted at similar principles in general well-posedness. The numerical example indicated the connection between well-posedness and financial analysis. In the last two chapters, we presented a summarized research regarding the nuances and applications of well-posedness.

For future consideration, I would be interested in a different approach to proof of Theorem 1.6.1. using the concept of Furi-Vignoli criterion, since all notions of well-posedness can be similarly characterized by level sets. The concepts of well-posed optimization problems span across all fields of science and are extensively researched as shown in the last chapter and I have barely scratched the surface.

References

1. Ferrentino, Rosa; Boniello, Carmine. On the Well-Posedness for Optimization Problems: A Theoretical Investigation. *Applied Mathematics*. 2019, vol. 10, no. 1, pp. 19–38.
2. Lucchetti, Roberto. *Convexity and well-posed problems*. Springer Science & Business Media, 2006.
3. Dontchev, Assen L; Zolezzi, Tullio. *Well-posed optimization problems*. Springer, 2006.
4. Revalski, Julian Petrov. Generic well-posedness in some classes of optimization problems. *Acta Universitatis Carolinae. Mathematica et Physica*. 1987, vol. 28, no. 2, pp. 117–125.
5. Revalski, JP; Zhivkov, NV. Well-posed constrained optimization problems in metric spaces. *Journal of optimization theory and applications*. 1993, vol. 76, pp. 145–163.
6. Kabanikhin, Sergei Igorevich. Definitions and examples of inverse and ill-posed problems. 2008.
7. Sizikov, Valery S et al. *Well-posed, ill-posed, and intermediate problems with applications*. Vol. 49. Walter de Gruyter, 2011.