Projection Algorithms and Monotone Operators

Heinz Bauschke

Dedicated to my mother
and
the memory of my father

Thesis Defence, CECM, SFU

Thursday, August 8, 1996, 10:30

Overview

1. Introduction to Part I

Suppose $C_1, \ldots, C_N$ are finitely many closed convex subsets of a Hilbert space X with

$$C := \bigcap_i C_i \neq \emptyset.$$

The *Convex Feasibility Problem* is simply:

(CFP) Find a point in C.

The sets C_i are referred to as the *constraints* and the set C is the set of all *solutions*.

Prototype: the constraints C_i are hyperplanes

$$\{x : \langle a_i, x\rangle = b_i\}$$

for some row vectors a_i of a matrix A and real components b_i of a vector b. Then

$$x \in C \Leftrightarrow Ax = b.$$

(Could also impose nonnegativity of solutions by adding the nonnegative orthant as a constraint.)

The CFP is very common in mathematical and physical sciences:

Best Approximation Theory
Constraints: closed subspaces.
Applications: statistics (linear prediction theory), complex analysis (Bergman kernels), partial differential equations (Dirichlet problem).

Discrete Image Reconstruction
Constraints: convex polyhedral sets; X is Euclidean.
Applications: medical imaging (computerized tomography), electron microscropy.

Subgradient algorithms
Constraints: sublevel sets of convex functions (approximated by supersets).
Applications: convex inequalities, minimization of convex (nonsmooth) functions.

Typically:
• one cannot find a solution in C directly, but
• each constraint set C_i is "simple" in the sense that its projection (nearest point mapping) is easy to compute.

Consequently, one tries to solve CFP algorithmically by generating a sequence of points that is supposed to converge to a solution.

The *(projection) algorithm* analyzed in Part I computes the next iterate from a current iterate x_n by

$$x_{n+1} := x_n + \alpha_n \rho_n \sum_{i=1}^{N} \omega_{i,n}(P_{i,n}x_n - x_n).$$

Here:
• $P_{i,n}$ is the *projection* onto some $C_{i,n} \supseteq C_i$;
• $\alpha_n \in [0, 2]$ is a *relaxation parameter*;
• ρ_n (≥ 1) is an *extrapolation parameter*;
• $\omega_{i,n} \geq 0$ are *weights*: $\sum_i \omega_{i,n} = 1$.

This framework is broad enough to cover MANY algorithms.

Important questions concerning sequences (x_n) generated by this algorithm are:

• when does (x_n) converge weakly to $x \in C$?
• when in norm?
• when linearly: $\|x_n - x\| = \mathcal{O}(\theta^n)$ for $\theta < 1$?

There are numerous apparently unrelated results for incarnations of the projection algorithm.

Some of my favourite incarnations of the projection algorithm follow. For simplicity, each $C_{i,n} \equiv C_i$. Denote the projection onto C_i, C by P_i, P, respectively. $x_0 \in X$ is the *starting point*.

Cyclic projections.

The sequence (x_n) is obtained by projecting cyclically onto the constraints:

$$(x_0, P_1x_0, P_2P_1x_0, \ldots, P_N \cdots P_1x_0,\\ P_1P_N \cdots P_1x_0, \ldots)$$

(That is: $\omega_{i,n} \in \{0, 1\}$, $\alpha_n \equiv 1$, $\rho_n \equiv 1$.)

For $N = 2$ constraints, we obtain the **method of alternating projections:**

$$(x_0, P_1x_0, P_2P_1x_0, P_1P_2P_1x_0, P_2P_1P_2P_1x_0, \ldots)$$

von Neumann/Halperin (1933/1962)
If each C_i is a closed subspace, then (x_n) converges in norm to Px_0.

Browder (1967)
If each C_i is a closed subspace with $C_1^\perp + \cdots + C_N^\perp$ closed, then (x_n) converges linearly to Px_0.

Kaczmarz (1937)
If each C_i is a hyperplane and X is Euclidean, then (x_n) converges to Px_0.

Gubin et al. (1967)
If $C_N \cap \bigcap_{i=1}^{N-1} \operatorname{int} C_i \neq \emptyset$, then (x_n) converges linearly to some point in C.

Bregman (1965)
(x_n) converges weakly to some point in C.

Remotest-set projections.

For a current iterate x_n, find first the most violated constraint, i.e., $i \in \{1, \ldots, N\}$ such that $\|x_n - P_i x_n\| = \max_j \|x_n - P_j x_n\|$ and then update

$$x_{n+1} := P_i x_n.$$

Agmon/Motzkin&Schoenberg (1954)
If each C_i is a halfspace, then (x_n) converges to some point in C.

Bregman (1965)
(x_n) converges weakly to some point in C.

Gubin et al. (1967)
If $C_N \cap \bigcap_{i=1}^{N-1} \operatorname{int} C_i \neq \emptyset$, then (x_n) converges linearly to some point in C.

Parallel projections.

Cimmino (1938)
Suppose each C_i is a hyperplane, X is finite-dimensional, $\omega_{i,n} \equiv 1/N$, $\alpha_n \equiv 2$, and $\rho_n \equiv 1$. Then (x_n) converges to some point in C.

Pierra (1984)
Suppose $\omega_{i,n} \equiv 1/N$, $\alpha_n \equiv \alpha \leq 1$, and

$$\rho_n := \begin{cases} \dfrac{\sum_i \omega_{i,n} \|x_n - P_i x_n\|^2}{\|\sum_i \omega_{i,n}(x_n - P_i x_n)\|^2}, & \text{if } x_n \notin \bigcap_{i \in I_n} C_{i,n}; \\ 1, & \text{otherwise.} \end{cases}$$

If X is Euclidean or $\bigcap_i \operatorname{int} C_i \neq \emptyset$, then (x_n) converges in norm to some point in C.

The aim of Part I is to analyze the projection algorithm in detail, to bring out underlying recurring key concepts, and to improve, unify, and review existing results.

2. Projections

We want to be able to compute the **projection** onto a closed convex nonempty set C which we denote by P_C or P.

Fact. Suppose $y \in X$. Then:
(i) The (unique) point Py is characterized by

$$Py \in C \quad \text{and} \quad \langle C - Py, y - Py \rangle \leq 0.$$

(ii) For every $x \in X$,

$$\begin{aligned} \|x - y\|^2 = \|Px - Py\|^2 & \\ + \|(I - P)x - (I - P)y\|^2 & \\ + 2 \underbrace{\langle x - Px, Px - Py \rangle}_{\geq 0} & \\ + 2 \underbrace{\langle y - Py, Py - Px \rangle}_{\geq 0} & . \end{aligned}$$

In particular, P is *(firmly) nonexpansive*.

Some explicit examples.

unit ball $C = \{x \in X : \|x\| \leq 1\}$. Then $P_C x = x$, if $x \in C$; $P_C x = x/\|x\|$, otherwise.

nonnegative orthant $C = \{x \in X : x \geq 0\}$. Then $P_C x = x^+$.

hyperplane $C = \{x \in X : \langle a, x\rangle = b\}$, where $a \neq 0$ and $b \in \mathbb{R}$. Then

$$P_C x = x - \frac{\langle a, x\rangle - b}{\|a\|^2} a.$$

halfspace $C = \{x \in X : \langle a, x\rangle \leq b\}$, where $a \neq 0$ and $b \in \mathbb{R}$. Then

$$P_C x = x - \frac{(\langle a, x\rangle - b)^+}{\|a\|^2} a.$$

subspace $C = \text{span}\{a_1, \ldots, a_n\}$, where a_i are linearly independent column vectors of a matrix A. Then $P_C x = A(A^*A)^{-1}A^*x$.

3. (bounded) (linear) regularity

For simplicity, we define these notions first for two closed convex sets C_1, C_2 in X with $C := C_1 \cap C_2 \neq \emptyset$.

We say that $\{C_1, C_2\}$ is ...

regular, if $\forall$ (x_n) in X:

$$\max\{d(x_n, C_1), d(x_n, C_2)\} \to 0 \quad \Rightarrow \quad d(x_n, C) \to 0.$$

boundedly regular, if $\forall$ bounded (x_n) in X:

$$\max\{d(x_n, C_1), d(x_n, C_2)\} \to 0 \quad \Rightarrow \quad d(x_n, C) \to 0.$$

linearly regular, if $\exists \kappa > 0$ such that

$$d(x, C) \leq \kappa \max\{d(x, C_1), d(x, C_2)\}, \quad \forall x \in X.$$

boundedly linearly regular, if $\forall$ bounded $S \subseteq X$, $\exists \kappa_S > 0$ such that

$$d(x, C) \leq \kappa_S \max\{d(x, C_1), d(x, C_2)\}, \quad \forall x \in S.$$

The following implications are immediate from the definitions:

$$\begin{array}{ccc} \text{linearly regular} & \Rightarrow & \text{boundedly linearly regular} \\ \Downarrow & & \Downarrow \\ \text{regular} & \Rightarrow & \text{boundedly regular.} \end{array}$$

If S is a convex subset of X, then the *strong relative interior* of S, denoted sri S, is defined by

$$x \in \text{sri}\, S \Leftrightarrow \text{cone}\,(S - x) = \overline{\text{span}}\,(S - x).$$

The best result on regularities is:

Fact. If $0 \in \text{sri}\,(C_1 - C_2)$, then $\{C_1, C_2\}$ is boundedly linearly regular.

Let's record some consequences and further facts.

Corollary. If $C_2 \cap \text{int}\, C_1 \neq \emptyset$, then $\{C_1, C_2\}$ is boundedly linearly regular.

Proof. $C_2 \cap \text{int}\, C_1 \neq \emptyset \Rightarrow 0 \in \text{sri}\,(C_1 - C_2)$. Ξ

Fact. Suppose C_1 and C_2 are closed subspaces. Then TFAE:
- $0 \in \text{sri}\,(C_1 - C_2)$.
- $C_1 + C_2$ is closed.
- $C_1^\perp + C_2^\perp$ is closed.
- The "angle" between C_1 and C_2 is positive.
- $\{C_1, C_2\}$ is (boundedly) (linearly) regular.

(This is false for cones.)

Fact. If X is finite-dimensional, then $\{C_1, C_2\}$ is boundedly regular.

Regularities for finitely many sets $C_1, \dots, C_N$ with $C := \bigcap_i C_i \neq \emptyset$ are defined analogously.

Some striking (and sharp) results are:

Fact. Suppose each C_i is a closed subspace. Then $\{C_1, \dots, C_N\}$ is (boundedly) (linearly) regular if and only if $C_1^\perp + \dots C_N^\perp$ is closed.

Fact. Suppose $C_N \cap \bigcap_{i=1}^{N-1} \operatorname{int} C_i \neq \emptyset$. Then $\{C_1, \dots, C_N\}$ is boundedly linearly regular.

Fact. (Hoffman; 1952) Suppose each C_i is a halfspace. Then $\{C_1, \dots, C_N\}$ is linearly regular.

Fact. Suppose $C_1, \dots, C_M$ are finitely many convex polyhedra, $C_{M+1}, \dots, C_N$ are finitely many closed convex sets, and X is Euclidean. If

$$\bigcap_{i=1}^{M} C_i \cap \bigcap_{j=M+1}^{N} \operatorname{ri} C_j \neq \emptyset,$$

then $\{C_1, \dots, C_N\}$ is boundedly linearly regular.

4. Fejér monotone sequences

Definition. A sequence (x_n) is **Fejér monotone** with respect to a closed convex nonempty set C in X, if

$$\|x_{n+1} - c\| \leq \|x_n - c\|, \qquad \forall n \in \mathbb{N}, c \in C.$$

Punchline: our sequences are!

Facts. Suppose (x_n) is Fejér monotone with respect to C. Then:
(i) The sequence $(P_C x_n)$ converges in norm, say $c^* := \lim_n P_C x_n \in C$.
(ii) (x_n) is bounded and $(d(x_n, C)) = (\|x_n - P_C x_n\|)$ is decreasing, hence convergent.
(iii) (x_n) is weakly convergent to c^* if and only if all weak cluster points of (x_n) lie in C.
(iv) (x_n) converges in norm to c^* if and only if $d(x_n, C) \to 0$.
(v) If there exists $\theta < 1$ such that $d(x_{n+1}, C) \leq \theta d(x_n, C)$, $\forall n$, then (x_n) converges linearly to c^* with rate θ.

5. A prototypical result

We now show how the key concepts (*projections*, *regularities*, *Fejér monotonicity*) work together by proving a prototypical result on the method of alternating projections:

$$\boxed{(x_0, P_1 x_0, P_2 P_1 x_0, P_1 P_2 P_1 x_0, P_2 P_1 P_2 P_1 x_0, \dots)}$$

Let (WLOG) $x := x_n$ and $y := P_i x_n$ be two consecutive iterates so that $i \in \{1,2\}$ and $x \in C_j$ for $\{j\} = \{1,2\} \setminus \{i\}$. Fix an arbitrary $c \in C = C_1 \cap C_2$. Then $\|y - c\| = \|P_i x - P_i c\| \leq \|x - c\|$; hence

$$\boxed{(x_n) \quad \text{is Fejér monotone with respect to C.}}$$

Also, $\|x - c\|^2 \geq \|P_i x - c\|^2 + \|x - P_i x\|^2$, which yields $d^2(x, C) \geq d^2(y, C) + d^2(x, C_i)$. Since $x \in C_j$, we obtain further $\max\{d^2(x, C_i), d^2(x, C_j)\} \leq d^2(x, C) - d^2(y, C)$, which translates back to

$$\boxed{\begin{aligned} & \max\{d^2(x_n, C_1), d^2(x_n, C_2)\} \\ \leq \quad & d^2(x_n, C) - d^2(x_{n+1}, C), \quad \forall n \geq 1. \end{aligned}}$$

Theorem. Suppose C_1, C_2 are closed convex subsets of X with $C := C_1 \cap C_2 \neq \emptyset$. Suppose further $(x_n)_{n \geq 0}$ is Fejér monotone with respect to C with

$$\begin{aligned} & \max\{d^2(x_n, C_1), d^2(x_n, C_2)\} \\ \leq \quad & d^2(x_n, C) - d^2(x_{n+1}, C), \quad \forall n. \end{aligned}$$

Let $c^* := \lim_n P_C x_n$. Then:
(i) (x_n) converges weakly to c^*.
(ii) If $\{C_1, C_2\}$ is boundedly regular, then (x_n) converges in norm to c^*.
(iii) If $\{C_1, C_2\}$ is boundedly linearly regular, then (x_n) converges linearly to c^*.
(iv) If $\{C_1, C_2\}$ is linearly regular, then (x_n) converges linearly to c^* with a rate independent of the starting point.

Remark. If (x_n) is a sequence of alternating projections, then the Theorem is applicable (see previous page).

Proof. The sequence (x_n) is bounded and so is $S := \{x_n : n \geq 0\}$. Also, the sequence $(d(x_n, C))$ is convergent; hence $d^2(x_n, C) - d^2(x_{n+1}, C) \to 0$, which yields

$$(*) \qquad \max\{d(x_n, C_1), d(x_n, C_2)\} \to 0.$$

(i): $(*)$ implies that all weak cluster points of (x_n) lie in C. Apply **Fejér Facts (iii)**.
(ii): Bounded regularity and $(*)$ yield $d(x_n, C) \to 0$, which is equivalent to $x_n \to c^*$, by **Fejér Facts (iv)**.
(iii): There exists $\kappa_S > 0$ such that $d(x_n, C) \leq \kappa_S \max\{d(x_n, C_1), d(x_n, C_2)\}$, $\forall n$. Hence $d^2(x_n, C) \leq \kappa_S^2(d^2(x_n, C) - d^2(x_{n+1}, C))$, $\forall n$, which implies that (x_n) converges linearly to c^* with rate $\sqrt{1 - 1/\kappa_S^2}$ (**Fejér Facts (v)**).
(iv): Analogous to **(iii)** with the difference that we can pick κ_S *independent* of S. $\quad \Xi$

Remark. If C is an affine subspace, then $c^* = P_C x_0$ and we obtain *best approximation* results.

6. Conclusion of Part I

We have improved, unified, and reviewed many existing results on projection algorithms by using the following key modules:

- projections and their properties;
- (bounded) (linear) regularity;
- Fejér monotone sequences.

The tools employed are from the beautiful and powerful area of *Convex Analysis*.

7. Introduction to Part II

First motivation. Suppose A is an $n \times n$ matrix that is positive semi-definite, i.e., $x^t A x = \langle x, Ax \rangle \geq 0$, $\forall x \in \mathbb{R}^n$. Then the transpose of A is also positive semi-definite.

Question: More generally, is this true for operators defined on Banach spaces??

Second motivation. Recently, several notions of monotonicity have been coined that:
- imply *maximal monotonicity*;
- are automatic in *reflexive* spaces;
- hold for *subdifferentials* of convex functions.

Question: What about continuous linear positive semi-definite operators??

Throughout: X is some real Banach space with dual space X^*. If $x^* \in X^*$ and $x \in X$, then $\langle x^*, x \rangle$ or $\langle x, x^* \rangle$ is x^* evaluated at x.

Definition. A set-valued map T from X to X^* is a monotone operator, if

$$\langle Tx - Ty, x - y \rangle \geq 0, \quad \forall x, y \in X.$$

T is maximal monotone, if T is monotone and the *graph* of T is maximal in $X \times X^*$ with respect to set-inclusion.

Example. (Rockafellar) The *subdifferential map* ∂f is maximal monotone, for every convex lower semi-continuous proper function f on X.

Example. Every *continuous linear positive semi-definite operator* on X is maximal monotone.

The zoo. (Gossez, Simons, Fitzpatrick&Phelps) Suppose T is maximal monotone on X. Define (set-valued) extensions of T whose graphs reside in $X^{**} \times X^*$ as follows:

$\boxed{x^* \in T_1 x^{**}}$, if $\exists$ bounded net $(x_\alpha, x_\alpha^*) \in \text{graph}(T)$ with $x_\alpha \overset{\text{w*}}{\to} x^{**}$ and $x_\alpha^* \to x^*$.

$\boxed{x^* \in T_0 x^{**}}$, if $\inf_{y \in X} \langle Ty - x^*, y - x^{**} \rangle = 0$.

$\boxed{x^* \in \overline{T} x^{**}}$, if $\inf_{y \in X} \langle Ty - x^*, y - x^{**} \rangle \geq 0$.

Then T is:

(D) or "dense", if $T_1 = \overline{T}$.

(RD) or "range-dense", if range $T_1 = \text{range}\,\overline{T}$.

(NI) or "nonnegative infimum", if $T_0 = \overline{T}$.

(LMM) or "locally maximal monotone", if $\forall$ weak* closed convex bounded subset C of X^* with range $T \cap \text{int}\, C \neq \emptyset$, and $\forall$ $x_0 \in X$, $x_0^* \in (\text{int}\, C) \setminus Tx_0$, $\exists$ $(z, z^*) \in \text{graph}(T) \cap (X \times C)$ with $\langle z^* - x_0^*, z - x_0 \rangle < 0$.

Facts. (Gossez and Simons)

- In general: (D) $\Rightarrow$ (RD) $\Rightarrow$ (NI).
- (D) and (LMM) hold in *reflexive* spaces.
- *Subdifferentials* are (D) and (LMM).

Question: What is (D), (RD), (NI), (LMM) for continuous linear monotone (a.k.a. positive semi-definite) operators??

The aim of Part II is to study the various monotonicities for continuous linear positive semi-definite operators.

7. The tools

Tool. (Decomposition/Quadratic Function) Suppose T is a continuous linear monotone operator on X. Then T can be written *uniquely* as the sum of two continuous linear operators, $T = P + S$, where P is *symmetric* (i.e. $P^*|_X = P$), and S is *skew* (i.e. $S^*|_X = -S$). In fact:

$$P := \frac{T + T^*|_X}{2} \quad \text{and} \quad S := \frac{T - T^*|_X}{2}.$$

Let $q(x) := \frac{1}{2}\langle x, Tx \rangle$, $\forall x \in X$. Then

$$\partial q(x) = \{\nabla q(x)\} = Px, \quad \forall x \in X;$$

hence P is (as the subdifferential of a continuous convex function) *extremely nice*; for instance, P is (D). Although S is far away from being a subdifferential, it has the good property that $\langle Sx, x \rangle = 0$, $\forall x \in X$.

Key Tool. (Fenchel's Duality Theorem) Suppose A is a continuous linear operator from X to some Banach space Y. Suppose further f is a convex lower semi-continuous proper function on X as is g on Y. Define

$$p := \inf_{x \in X} \{f(x) + g(Ax)\}$$

and

$$d := -\inf_{y^* \in Y^*} \{f^*(-A^*y^*) + g^*(y^*)\}.$$

Then $p \geq d$. If $A(\text{dom}\, f) \cap \text{int}\,\text{dom}\, g \neq \emptyset$ and p is finite, then $p = d$ and d is attained.

Reminder: The Fenchel conjugate f^* of f is defined by

$$f^*(x^*) := \sup_{x \in X} \langle x^*, x\rangle - f(x),\ \forall x^* \in X^*.$$

8. The main results

Theorem. Suppose T is a continuous linear monotone operator on X with skew part S. Then TFAE:

(i) T is (D).
(ii) T is (RD).
(iii) T is (NI).
(iv) T is (LMM).
(v) T^* is monotone.

(vi) S is (D).
(vii) S is (RD).
(viii) S is (NI).
(ix) S is (LMM).
(x) S^* is monotone.

Remark. "(v)⇒(i)" gives an affirmative answers to an old question by Gossez.

Theorem. Suppose T is a continuous linear monotone operator on X with skew part S. Then TFAE:

(i) T and $T^*|_X$ are (D).
(ii) T and $T^*|_X$ are (NI).
(iii) T and $T^*|_X$ are (LMM).
(iv) T^* and $(T^*|_X)^*$ are monotone.
(v) T is weakly compact.

(vi) S and $-S$ are (D).
(vii) S and $-S$ are (NI).
(viii) S and $-S$ are (LMM).
(ix) S^* and $-S^*$ are monotone.
(x) S^* is skew.
(xi) S is weakly compact.

Remark. We can interpret monotonicity of T^* as "one half" of weak compactness of T.

9. "Weird" examples and (cms) spaces

If T is a continuous linear monotone operator on X with skew part S, then T will satisfy one of the following alternatives:

"**good**": S^* and $-S^*$ are both monotone.
"**so-so**": only one of $\{S^*, -S^*\}$ is monotone.
"**bad**": neither S^* nor $-S^*$ is monotone.

Question: Do these cases all happen?

Gossez's Example. Let G from ℓ_1 to $\ell_\infty = \ell_1^*$ be given by the infinite matrix:

$$\begin{pmatrix} 0 & 1 & 1 & 1 & \cdots \\ -1 & 0 & 1 & 1 & \cdots \\ -1 & -1 & 0 & 1 & \cdots \\ -1 & -1 & -1 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Then G is "so-so":
G is skew,
G^* is not monotone, but
$-G^*$ is monotone.

Note that

$$(Gx)_n := -\sum_{k<n} x_k + \sum_{k>n} x_k, \ \forall x \in \ell_1 \ \forall n.$$

This suggests a "continuous" version.

Fitzpatrick&Phelps's Example. Define F from $L_1[0,1]$ to $L_\infty[0,1] = L_1^*[0,1]$ by

$$(Fx)(t) := -\int_0^t x(s)\,ds + \int_t^1 x(s)\,ds, \ \forall x \in L_1[0,1] \ \forall t.$$

Then F is "bad":
F is skew,
F^* is not monotone, and
$-F^*$ is not monotone.

Remark. I re-derived these examples systematically and with less pain.

Definition. The Banach space X is (cms) or "a conjugate monotone space", if the conjugate of every continuous linear monotone operator on X is again monotone.
The Theorems yield: X is (cms) if and only if *every continuous linear monotone operator on X is "good" (or weakly compact).*

Some (cms) Banach spaces:
reflexive spaces; in particular, ℓ_p and $L_p[0,1]$ for $1 < p < \infty$, and Hilbert spaces.

Some Banach spaces that are not (cms):
ℓ_1, $L_1[0,1]$, and their biduals; every space that contains a complemented copy of ℓ_1. (Lift the "so-so" and "bad" examples!)

(cms) Banach lattices:
are *precisely* those that do not contain a complemented copy of ℓ_1. (Uses deeper Banach Space Theory.) In particular: c_0, c, ℓ_∞, $L_\infty[0,1]$, and $C[0,1]$ are (cms).

We now know that the "monotonicities" (D), (RD), (NI), and (LMM) all coincide
- for subdifferentials;
- in reflexive spaces;
- *for continuous linear monotone operators.*

Question: Can they actually differ??

Still open, next best candidates for counterexamples are *regularizations*, i.e., maps of the form

$$T + \lambda J,$$

where T is continuous linear monotone on X, $\lambda > 0$, and $J := \partial \frac{1}{2}\|\cdot\|^2$ is the *duality map*.

Theorem. TFAE in c_0, c, ℓ_1, ℓ_∞, $L_1[0,1]$, $L_\infty[0,1]$, and $C[0,1]$:
- T is (D).
- $T + \lambda J$ is (RD), $\forall \lambda \geq 0$.
- $T + \lambda J$ is (LMM), $\forall \lambda \geq 0$.

11. Conclusion of Part II

We have shown that the various monotonicities all coincide for continuous linear monotone operators although they do not hold automatically.

The study depended on results from Functional Analysis, Convex Analysis, and Banach Space Theory, but most importantly on

- Fenchel's Duality Theorem,

which continues to amaze me by its wide range of applications.

The End