Functional Analysis and Optimization

Kazufumi Ito*

November 4, 2014

Abstract

In this monograph we develop the function space method for optimization problems and operator equations in Banach spaces. Optimization is the one of key components for mathematical modeling of real world problems and the solution method provides an accurate and essential description and validation of the mathematical model. Optimization problems are encountered frequently in engineering and sciences and have widespread practical applications. In general we optimize an appropriately chosen cost functional subject to constraints. For example, the constraints are in the form of equality constraints and inequality constraints. The problem is casted in a function space and it is important to formulate the problem in a proper function space framework in order to develop the accurate theory and the effective algorithm. Nonsmooth optimization becomes a very basic modeling toll and enlarges and enhances the applications of the optimization method in general. For example, in the classical variational formulation we analyze the non Newtonian energy functional and nonsmooth friction penalty. For the imaging/signal analysis the sparsity optimization is used by means of of L^1 and TV regularization. In order to develop an efficient solution method for large scale optimizations it is essential to develop a set of necessary conditions in the equation form rather than the variational inequality. The Lagrange multiplier theory for the constrained minimizations and nonsmooth optimization problems is developed and analyzed The theory derives the complementarity condition for the multiplier and the solution. Consequently, one can derive the necessary optimality system for the solution and the multiplier in a general class of nonsmooth and noncovex optimizations. In the light of the theory we derive and analyze numerical algorithms based the primal-dual active set method and the semi-smooth Newton method.

The monograph also covers the basic materials for real analysis, functional analysis, Banach space theory, convex analysis, operator theory and PDE theory, which makes the book self-contained and comprehensive. The analysis component is naturally connected to the optimization theory. The necessary optimality condition is in general written as nonlinear operator equations for the primal variable and Lagrange multiplier. The Lagrange multiplier theory of a general class of nonsmooth and non-convex optimization are based on the functional analysis tool. For example, the necessary optimality for the constrained optimization is in general nonsmooth and a psuedo-monontone type operator equation. To develop the mathematical treatment

*Department of Mathematics, North Carolina State University, Raleigh, North Carolina, USA

of the PDE constrained minimization we also present the PDE theory and linear and nonlinear C_0-semigroup theory on Banach spaces for well-posedness of the control dynamics and optimization problems. That is, the existence and uniqueness of solutions to a general class of controlled linear and nonlinear PDEs and Cauchy problems for nonlinear evolution equations in Banach space are discussed. Especially the linear operator theory and the (psuedo-) monotone operator and mapping theory and the fixed point theory. Throughout the monograph demonstrates the theory and algorithm using concrete examples and describes how to apply it for motivated applications

1 Introduction

In this monograph we develop the function space approach for the optimization problems, e.g., nonlinear programming in Banach spaces, convex and non-convex nonsmooth variational problems, control and inverse problems, image/signal analysis, material design, classification, and resource allocation and for the nonlinear equations and Cauchy problems in Banach spaces.

A general class of constrained optimization problems in function spaces is considered and we develop the Lagrange multiplier theory and effective solution algorithms. Applications are presented and analyzed for various examples including control and design optimization, inverse problems, image analysis and variational problems. Non-convex and non-smooth optimization becomes increasing important for advanced and effective use of the optimization methods and also the essential topic for the monograph We introduce the basic function space and functional analysis tools needed for our analysis. The monograph provides an introduction to the functional analysis, real analysis and convex analysis and their basic concepts with illustrated examples. Thus, one can use the book as a basic course material for the functional analysis and nonlinear operator theory. An emerging application of optimization is the imaging analysis and can be formulated as the PDE's constrained optimization. The PDE theory and the nonlinear operator theory and their applications are analyzed in details.

There are numerous applications of the optimization theory and variational method in all sciences and to real world applications. For example, the followings are the motivated examples.

Example 1 (Quadratic programming)

$$\min \quad \tfrac{1}{2}x^t Ax - b^t x \text{ over } x \in R^n$$

$$\text{subject to} \quad Ex = c \text{ (equality constraint)} \quad Gx \leq g \text{ (inequality constraint).}$$

where $A \in R^{n\times n}$ is a symmetric positive matrix, $E \in R^{m\times n}$, $G \in R^{p\times n}$, and $b \in R^n$, $c \in R^m$, $g \in R^p$ are given vectors. The quadratic programming is formulated on a Hilbert space X for the variational problem, in which $(x, Ax)_X$ defines the quadratic form on X.

Example 2 (Linear programming) The linear programing minimizes $(c, x)_X$ subject to the linear constraint $Ax = b$ and $x \geq 0$. Its dual problem is to maximize (b, x) subject to $A^*x \leq c$. It has many important classes of applications. For example, the optimal transport problem is to find the joint distribution function $\mu(x, y) \geq 0$ on $X \times Y$ that transports the probability density $p_0(x)$ to $p_1(y)$, i.e.

$$\int_Y \mu(x,y)\,dx = p_0(x), \quad \int_Y \mu(x,y)\,dy = p_1(y),$$

with minimum energy

$$\int_{X\times Y} c(x,y)\mu(x,y)\,dxdy = (c,\mu).$$

Example 3 (Integer programming) Let F is a continuous function R^n. The integer programming is to minimize $F(x)$ over the integer variables x, i.e.,

$$\min \quad F(x) \quad \text{subject to} \quad x \in \{0,1,\cdots N\}.$$

The binary optimization is to minimize $F(x)$ over the binary variables x. The network optimization problem is the one of important applications.

Example4 (Obstacle problem) Consider the variational problem;

$$\min \quad J(u) = \int_0^1 (\frac{1}{2}|\frac{d}{dx}u(x)|^2 - f(x)u(x))\,dx \tag{1.1}$$

subject to

$$u(x) \le \psi(x)$$

over all displacement function $u(x) \in H_0^1(0,1)$, where the function ψ represents the upper bound (obstacle) of the deformation. The functional J represents the Newtonian deformation energy and $f(x)$ is the applied force.

Example 5 (Function Interpolation) Consider the interpolation problem;

$$\min \quad \int_0^1 (\frac{1}{2}|\frac{d^2}{dx^2}u(x)|^2\,dx \quad \text{over all functions } u(x), \tag{1.2}$$

subject to the interpolation conditions

$$u(x_i) = b_i, \quad \frac{d}{dx}u(x_i) = c_i,\ 1 \le i \le m.$$

Example 6 (Image/Signal Analysis) Consider the deconvolution problem of finding the image u that satisfies $(Ku)(x) = f(x)$ for the convolution K

$$(Ku)(x) = \int_\Omega k(x,y)u(y)\,dy.$$

over a domain Ω. It is in general ill-posed and we use the (Tikhonov) regularization formulation to have a stable deconvolution solution u, i.e., for $\Omega = [0,1]$

$$\min \quad \int_0^1 (|Ku-f|^2 + \frac{\alpha}{2}|\frac{d}{dx}u(x)|^2 + \beta\,|u(x)|)\,dx \text{ over all possible image } u \ge 0, \tag{1.3}$$

where $\alpha,\ \beta \ge 0$ are the regularization parameters and the two regularization functional is chosen so that the variation of image u is regularized.

Example 7 (L^0 optimization) Problem of optimal resource allocation and material distribution and optimal time scheduling can be formulated as

$$F(u) + \int_\Omega h(u(\omega))\, d\omega \tag{1.4}$$

with $h(u) = \frac{\alpha}{2}|u|^2 + \beta\,|u|^0$ where

$$|0|^0 = 0, \quad |u|^0 = 1, \ \ otherwise.$$

Since

$$\int_\Omega |u(\omega)|^0\, d\omega = \mathrm{vol}(\{u(\omega) \neq 0\}),$$

the solution to (1.4) provides the optimal distribution of u for minimizing the performance $F(u)$ with appropriate choice of (regularization) parameters $\alpha,\ \beta > 0$.

Example 8 (Bayesian statistics) Consider the statical optimization;

$$\min \quad -\log\left(p(y|x) + \beta p(x)\right) \text{ over all admissible parameters } x, \tag{1.5}$$

where $p(y|x)$ is the conditional probability density of observation y given parameter x and the $p(x)$ is the prior probability density for parameter x. It is the maximum posterior estimation of the parameter x given observation y. In the case of inverse problems it has the form

$$\phi(x|y) + \alpha\, p(x) \tag{1.6}$$

for x is in a function space X and $\phi(x|y)$ is the fidelity and p is a regularization semi-norm on X. A constant $\alpha > 0$ is the Tikhonov regularization parameter.

Multi-dimensional versions of these examples and more illustrated examples will be discussed and analyzed throughout the monograph.

Discretization An important aspect of the function space optimization is the approximation method and theory. For example, let $B_i^n(x)$ is the linear B-spline defined by

$$B_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}} & \text{on } [x_{i-1}, x_i] \\ \frac{x-x_{i+1}}{x_i-x_{i+1}} & \text{on } [x_i, x_{i+1}] \\ 0 & \text{otherwise} \end{cases} \tag{1.7}$$

and let

$$u^n(x) = \sum_{i=1}^{n-1} u_i B_i(x) \sim u(x). \tag{1.8}$$

Then,

$$\frac{d}{dx}u^n = \frac{u_i - u_{i-1}}{x_i - x_{i-1}}, \quad x \in (x_{i-1}, x_i).$$

Let $x_i = \frac{i}{n}$ and $\Delta x = \frac{1}{n}$ and one can define the discretized problem of Example 1:

$$\min \quad \sum_{i=1}^{n} \left(\frac{1}{2}\left|\frac{u_i - u_{i-1}}{\Delta x}\right|^2 - f_i u_i\right) \Delta x$$

subject to

$$u_0 = u_n = 0, \quad u_i \le \psi_i = \psi(x_i).$$

where $f_i = f(x_i)$. The discretized problem is the quadratic programming for $\vec{u} \in R^{n-1}$ with

$$A = \frac{1}{\Delta x} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad b = \Delta x \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-2} \\ f_{n-1} \end{pmatrix}$$

Thus, we will discuss a convergence analysis of the approximate solutions $\{u^n\}$ to the solution u as $n \to \infty$ in an appropriate function space.

In general, one has the optimization of the form

$$\min \quad F_0(x) + F_1(x) \text{ over } x \in X \quad \text{subject to } E(x) \in \mathcal{K}$$

where X is a Banach space, F_0 is C^1 and F_1 is nonsmooth functional and $E(x) \in \mathcal{K}$ represents the constraint conditions with $\mathcal{K}$ is a closed convex cone. For the existence of a minimizer we require a coercivity condition and lower-semicontinuity and the continuity of the constraint in a weak or weak star topology of X. For the necessary optimality condition, we develop a generalized Lagrange multiplier theory:

$$\begin{cases} F'(x^*) + \lambda + E'(x^*)^*\mu = 0, \quad E(x) \in \mathcal{K} \\ \\ C_1(x^*, \lambda) = 0 \\ \\ C_2(x^*, \mu) = 0, \end{cases} \tag{1.9}$$

where λ is the Lagrange multiplier for a relaxed derivative of F_1 and μ is the Lagrange multiplier for the constraint $E(x) \in \mathcal{K}$. We derive the complementarity conditions C_1 and C_2 so that (1.9) is a complete set of equation for determining (x^*, λ, μ).

We discuss the solution method for linear and nonlinear equations in Banach spaces of the form

$$f \in A(x) \tag{1.10}$$

where $f \in X^*$ in the dual space of a Banach space and A is pseudo-monotone mapping $X \to X^*$. and the Cauchy problem:

$$\frac{d}{dt}x(t) \in A(t)(x(t)) + f(t), \quad x(0) = x_0, \tag{1.11}$$

where $A(t)$ is an m-dissipative mapping in $X \times X$. Also, we note that the necessary optimality (1.9) is a nonlinear operator equation and we use the nonlinear operator theory to analyze solutions to (1.9). We also discuss the PDE constrained optimization problems, in which f and $f(t)$ are a function of control and design and medium parameters. We minimize a proper performance index defined for the state and control functions subject to the constraint (1.10) or (1.11). Thus, we combine the analysis of the PDE theory and the optimization theory to develop a comprehensive treatment and analysis of the PDE constrained optimization problems.

The outline of the monograph is as follows. In Section 2 we introduce the basic function space and functional analysis tools needed for our analysis. It provides the introduction to the functional analysis and real analysis and their basic concepts with illustrated examples. For example, the dual space of normed space, Hahn-Banach theorem, open mapping theory, closed range theory, distribution theory and weak solution to PDEs, compact operator and spectral theory.

In Section 3 we develop the Lagrange multiplier method for a general class of (smooth) constrained optimization problems in Banach spaces. The Hilbert space theory is introduced first and the Lagrange multiplier is defined as the limit of the penalized problem. The constraint qualification condition is introduced and we derive the (normal) necessary optimality system for the primal and dual variables. We discuss the minimum norm problem and the duality principle in Banach spaces. In Section 4 we present the linear operator theory and C_0-semigroup theory on Banach spaces. In Section 5 we develop the monotone operator theory and pseudo-monotone operator theory for nonlinear equations in Banach spaces and the nonlinear semigroup theory in Banach spaces. The fixed point theory for nonlinear equations in Banach spaces.

In Section 7 we develop the basic convex analysis tools and the Lagrange multiplier theory for a general class of convex non-smooth optimization. We introduce and analyze the augmented Lagrangian method for the constrained optimization and non-smooth optimization problems.

In Section 8 we develop the Lagrange multiplier theory for a general class of non-smooth and non-convex optimizations.

2 Basic Function Space Theory

In this section we discuss the basic theory and concept of the functional analysis and introduce the function spaces, which is essential to our function space analysis of a general class of optimizations. We refer to the more comprehensive treatise and discussion on the functional analysis for e.g., Yosida [].

2.1 Complete Normed Space

In this section we introduce the vector space, normed space and Banach spaces.

Definition (Vector Space) A vector space over the scalar field K is a set X, whose elements are called vectors, and in which two operations, *addition* and *scalar multiplication*, are defined, with the following familiar algebraic properties.

(a) To every pair of vectors x and y corresponds a vector $x+y$, in such a way that

$$x+y=y+x \quad \text{and} \quad x+(y+z)=(x+y)+z;$$

X contains a unique vector 0 (the zero vector) such that $x+0=0+x$ for every $x \in X$, and to each $x \in X$ corresponds a unique *inverse* vector $-x$ such that $x+(-x)=0$.

(b) To every pair (α, x) with $\alpha \in K$ and $x \in X$ corresponds a vector $\alpha\, x$, in such a way that

$$1\, x = x, \quad \alpha(\beta x) = (\alpha\beta)\, x,$$

and such that the two distributive laws

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

hold.

The symbol 0 will of course also be used for the zero element of the scalar field. It is easy to see that the inverse element $-x$ of the additivity satisfies $-x = (-1)\,x$ since $0\,x = (1-1)\,x = x + (-1)\,x$ for each $x \in X$. In what follows we consider vector spaces only over the real number field R or the complex number field C.

Definition (Subspace, Basis) (1) A subset S of X is a (linear) subspace of a vector space X if $\alpha\,x_1 + \beta\,x_2 \in S$ for all $x_, x_2 \in S$ and and $\alpha,\ \beta \in K$.
(2) A family of linear vectors $\{x_i\}_{i=1}^n$ are linearly independent if $\sum_{i=1}^n \alpha_i\,x_i = 0$ if and only if $\alpha_i = 0$ for all $1 \le i \le n$. The dimension $dim(X)$ is the largest number of independent vectors in X.
(3) A basis $\{e_\alpha\}$ of a vector space X is a set of linearly independent vectors that, in a linear combination, can represent every vector in X, i.e., $x = \sum_\alpha a_\alpha e_\alpha$.

Example (Vector Space) (1) Let $X = C[0,1]$ be the vector space of continuous functions f on $[0,1]$. X is a vector space by

$$(\alpha\,f + \beta\,g)(t) = \alpha\,f(t) + \beta\,g(t), \quad t \in [0,1]$$

S= a space of polynomials of order less than n is a linear subspace of X and $dim(X) = n$. A family $\{t^k\}_{k=0}^\infty$ of polynomials is a basis of X and $dom(X) = \infty$.
(2) The vector space of real number sequences $x = (x_1, x_2, x_3, \cdots)$ is a vector space by

$$(\alpha\,x + \beta\,y)_k = \alpha\,x_k + \beta\,y_k, \quad k \in N.$$

A family of unit vectors $e^k = (0, \cdots, 1, 0, \cdots),\ k \in N$ is a basis.

(3) Let X be the vector space of square integrable functions on $(0,1)$. Let $e_k(x) = \sqrt(2)\sin(k\pi x),\ 0 \le x \le 1$. Then $\{e_k\}$ are orthonomal basis, i.e. $(e_k, e_j) = \displaystyle\int_0^1 e_k(x)e_j(x)\,dx = \delta_{k,j}$ and

$$f(x) = \sum_{k=1}^\infty f_k\,e_k(x), \text{ the Fourier sine series}$$

where $f_k = \int_0^1 f(x)e_k(x)\,dx$. e_{n+1} is not linearly dependent with $(e_1, \cdots, e_n)$. If so, $e_{n+1} = \sum_{i=1}^n \beta_i e_i$ but $\beta_i = (e_{n+1}, e_i) = 0$, which is the contradiction.

Definition (Normed Space) A norm $|\cdot|$ on a vector space X is a real-valued functional on X which satisfies

$$|x| \ge 0 \text{ for all } x \in X \text{ with equality if and only if } x = 0,$$

$$|\alpha\,x| = |\alpha|\,|x| \text{ for every } x \in x \text{ and } \alpha \in K$$

$$|x+y| \le |x| + |y| \quad \text{(triangle inequality) for every } x, y \in X.$$

A normed space is a vector space X which is equipped with a norm $|\cdot|$ and will be denoted by $(X, |\cdot|)$. The norm of a normed space X will be denoted by $|\cdot|_X$ or simply by $|\cdot|$ whenever the underlined space X is understood from the context.

Example (Normed space) Let $X = C[0,1]$ be the vector space of continuous functions f on $[0,1]$. The followings are examples the two normed space for X; X_1 is equipped with sup norm

$$|f|_{X_1} = \max_{x \in [0,1]} |f(x)|.$$

and X_2 is equipped with L^2 norm

$$|f|_{X_2} = \sqrt{\int_0^1 |f(x)|^2\,dx}.$$

<u>**Definition (Banach Space)**</u> Let X be a normed space.
(1) A sequence $\{x_n\}$ in a normed space X converges to the limit $x^* \in X$ if and only if $\lim_{n\to\infty} |x_n - x^*| = 0$ in R.
(2) A subspace S of a normed space X is said to be *dense* in X if each $x \in X$ is the limit of a sequence of elements in S.
(3) A sequence $\{x_n\}$ in a normed space X is called a Cauchy sequence if and only if $\lim_{m,\,n\to\infty} |x_m - x_n| = 0$, i.e., for all $\epsilon > 0$ there exists N such that for $m,\ n \geq N$ such that $|x_m - x_n| < \epsilon$. If every Cauchy sequence in X converges to a limit in X, then X is *complete* and X is called a *Banach* space.

A Cauchy sequence $\{x_n\}$ in a normed space X has a unique limit if it has accumulation points. In fact, for any limit x^*

$$\lim_{n\to\infty} |x^* - x_n| = \lim_{n\to\infty}\lim_{m\to\infty} |x_m - x_n| = 0.$$

<u>Example (continued)</u> Two norms $|x|_1$ and $|x|_2$ on a vector space X are equivalent if there exist $0 < \underline{c} \leq \bar{c} < \infty$ such that

$$\underline{c}|x|_1 \leq |x|_2 \leq \bar{c}\,|x|_1 \quad \text{for all} \;\; x \in X.$$

(1) All norms on R^n is are equivalent. But for $X = C[0,1]$ consider $x^n(t) = t^n$. Then, $|x^n|_{X_1} = 1$ but $|x^n|_{X_2} = \sqrt{\frac{1}{2n+1}}$ for all n. Thus, $|x^n|_{X_2} \to 0$ as $n \to \infty$ and the two norms X_1 and X_2 are not equivalent.
(2) Every bounded sequence in R^n has a convergent subsequence in R^n (Bolzano-Weierstrass theorem). Consider the sequence $x_n(t) = t^n$ in $X = C[0,1]$. Then x_n is not convergent despite $|x_n|_{X_1} = 1$ is bounded.

<u>**Definition (Hilbert Space)**</u> If X is a vector space, a functional $(\cdot,\cdot)$ defined on $X \times X$ is called an *inner product* on X provided that for every $x, y, z \in X$ and $\alpha, \beta \in C$

$$(x, y) = \overline{(y, x)}$$

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

$$(x, x) \geq 0 \text{ and } (x, x) = 0 \text{ if and only if } x = 0.$$

where $\bar{c}$ denotes the complex conjugate of $c \in C$. Given such an inner product, a norm on X can be defined by

$$|x| = (x, x)^{\frac{1}{2}}.$$

A vector space $(X, (\cdot,\cdot))$ equipped with an inner product is called a *pre–Hilbert space*. If X is complete with respect to the induced norm, then it is called a *Hilbert* space.

Every inner product satisfies the Cauchy-Schwarz inequality

$$|(x, y)| \leq |x|\,|y| \quad \text{for } x, y \in X$$

In fact, for all $t \in R$

$$|x + t(x,y)\,y|^2 = |x|^2 + 2t\,|(x,y)|^2 + t^2|(x,y)|^2\,|y|^2 \geq 0.$$

Thus, we must have $|(x,y)|^4 - |x|^2|(x,y)|^2|y|^2 \leq 0$ which shows the Cauchy-Schwarz inequality.

<u>Example (Hilbert space)</u> (1) R^n with inner product

$$(x,y) = \sum_{k=1}^{n} x_k y_k \quad \text{for } x, y \in R^n$$

(2) Consider a space ℓ^2 of square summable real number sequences $x = (x_1, x_2, \cdots)$ define the inner product

$$(x,y) = \sum_{k=1}^{\infty} x_k y_k \quad \text{for } x, y \in \ell^2$$

Then, ℓ^2 is an Hilbert space. In fact, let $\{x^n\}$ is a Cauchy sequence in ℓ^2. Since

$$|x^n_k - x^m_k| \leq |x^n - x^m|_{\ell^2},$$

$\{x^n_k\}$ is a Cauchy sequence in R for every k. Thus, one can define a sequence x by $x_k = \lim_{n\to\infty} x^n_k$. Since

$$\lim_{m\to\infty} |x^m - x^n|_{\ell^2} = |x - x_n|_{\ell^2}$$

for a fixed n, $x \in \ell^2$ and $|x_n - x|_{\ell^2} \to$ as $n \to \infty$.
(3) Consider a space X_3 of continuously differentiable functions on $[0,1]$ vanishing at $x = 0,\ x = 1$ with inner product

$$(f,g) = \int_0^1 \frac{d}{dx} f \frac{d}{dx} g\,dx \quad \text{for } \ f, g \in X_3.$$

Then X_3 is a Pre-Hilbert space.

Every normed space X is either a Banach space or a dense subspace of a Banach space Y whose norm $|x|_Y$ satisfies

$$|x|_Y = |x| \quad \text{for every } x \in X.$$

In this latter case Y is called the completion of X. The completion of a normed space proceeds as in Cantor's Construction of real numbers from rational numbers.

<u>**Completion**</u> The set of Cauchy sequences $\{x_n\}$ of the normed space X can be classified according to the equivalence $\{x_n\} \sim \{y_n\}$ if $\lim_{\,n\to\infty} |x_n - y_n| = 0$. Since $||x_n| - |x_m|| \leq |x_n - x_m|$ and R is complete, it follows that $\lim_{n\to\infty} |x_n|$ exists. We denote by $\{x_n\}'$, the class containing $\{x_n\}$. Then the set Y of all such equivalent classes $\tilde{x} = \{x_n\}'$ is a vector space by

$$\{x_n\}' + \{y_n\}' = \{x_n + y_n\}', \quad \alpha\,\{x_n\}' = \{\alpha\,x_n\}'$$

and we define

$$|\{x_n\}'|_Y = \lim_{n\to\infty} |x_n|.$$

It is easy to see that these definitions of the vector sum and scalar multiplication and norm do not depend on the particular representations for the classes $\{x_n\}'$ and $\{y_n\}'$. For example, if $\{x_m\} \sim \{y_n\}$, then $\lim_{\,n\to\infty} ||x_n| - |y_n|| \le \lim_{\,n\to\infty} |x_n - y_n| = 0$, and thus $\lim_{\,n\to\infty} |x_n| = \lim_{\,n\to\infty} |y_n|$. Since $|\{x_n\}'|_Y = 0$ implies that $\lim_{n\to\infty} |x_n - 0|$, we have $\{0\} \in \{x_n\}'$ and thus $|\cdot|_Y$ is a norm. Since an element $x \in X$ corresponds to the Cauchy sequence:

$$\bar{x} = \{x, x, \dots, x, \dots\}' \in Y,$$

X is naturally contained in Y.

To prove the completeness of Y, let $\{\tilde{x}_k\} = \{\{x_n^{(k)}\}\}$ be a Cauchy sequence in Y. Then for each k, we can choose an integer n_k such that

$$|x_m^{(k)} - x_{n_k}^{(k)}|_X \le k^{-1} \quad \text{for} \;\; m > n_k.$$

We show that the sequence $\{\tilde{x}_k\}$ converges to the class containing the Cauchy sequence $\{x_{n_k}^{(k)}\}$ of X.

To this end, note that

$$|\tilde{x}_k - \overline{x_{n_k}^{(k)}}|_Y = \lim_{m\to\infty} |x_m^{(k)} - x_{n_k}^{(k)}|_X \le k^{-1}.$$

Since

$$\begin{aligned} |x_{n_k}^{(k)} - x_{n_m}^{(m)}|_X &= |\overline{x_{n_k}^{(k)}} - \overline{x_{n_m}^{(m)}}|_Y \le |\overline{x_{n_k}^{(k)}} - \tilde{x}_k|_Y + |\tilde{x}_k - \tilde{x}_m|_Y + |\tilde{x}_m - \overline{x_{n_m}^{(m)}}|_Y \\ &\le k^{-1} + m^{-1} + |\tilde{x}_k - \tilde{x}_m|_Y, \end{aligned}$$

$\{x_{n_k}^{(k)}\}$ is a Cauchy sequence of X. Let $\tilde{x} = \{x_{n_k}^{(k)}\}'$. Then,

$$|\tilde{x} - \tilde{x}_k|_Y \le |\tilde{x} - \overline{x_{n_k}^{(k)}}|_Y + |\overline{x_{n_k}^{(k)}} - \tilde{x}_k|_Y \le |\tilde{x} - \overline{x_{n_k}^{(k)}}|_Y + k^{-1}.$$

Since, as shown above

$$|\tilde{x} - \overline{x_{n_k}^{(k)}}|_Y = \lim_{m\to\infty} |x_{n_m}^{m} - x_{n_k}^{(k)}|_X \le \lim_{m\to\infty} |\tilde{x}_m - \tilde{x}_k|_Y + k^{-1}$$

thus we prove that $\lim_{k\to\infty} |\tilde{x} - \overline{x_{n_k}^{(k)}}|_Y = 0$ and so $\lim_{k\to\infty} |\tilde{x} - \tilde{x}_k|_Y = 0$. Moreover, the above proof shows that the correspondence $x \in X$ to $\bar{x} \in Y$ is isomorphic and isometric and the image of X by this correspondence is dense in Y, that is X is dense in Y. $\square$

Example (Completion) Let $X = C[0,1]$ be the vector space of continuous functions f on $[0,1]$. We consider the two normed space for X; X_1 is equipped with sup norm

$$|f|_{X_1} = \max_{x\in[0,1]} |f(x)|.$$

and X_2 is equipped with L^2 norm

$$|f|_{X_2} = \sqrt{\int_0^1 |f(x)|^2 \, dx}$$

We claim that X_1 is complete but X_2 is not. The completion of X_2 is $L^2(0,1)$ =the space of square integrable functions.

First, discuss X_1. Let $\{f_n\}$ is a Cauchy sequence in X_1. Since

$$|f_n(x) - f_m(x)| \le |f_n - f_m|_{X_1},$$

$\{f_n(x)\}$ is Cauchy sequence in R for every $x \in [0,1]$. Thus, one can define a function f by $f(x) = \lim_{n\to\infty} f_n(x)$. Since

$$|f(x) - f(\hat{x})| \le |f_n(x) - f_n(\hat{x})| + |f_n(x) - f(x)| + |f_n(\hat{x}) - f(\hat{x})|$$

and $|f_n(x) - f(x)| \to 0$ uniformly $x \in [0,1]$, the candidate f is continuous. Thus, X_1 is complete.

Next, discuss X_2. Define a Cauchy sequence $\{f_n\}$ by

$$f_n(x) = \begin{cases} \frac{1}{2} + n\,(x - \frac{1}{2}) & \text{on } [\frac{1}{2} - \frac{1}{n}, \frac{1}{2} + \frac{1}{n}] \\ 0 & \text{on } [0, \frac{1}{2} - \frac{1}{n}] \\ 1 & \text{on } [\frac{1}{2} + \frac{1}{n}, 1] \end{cases}$$

The candidate limit f in this case is defined by $f(x) = 0,\ x < \frac{1}{2},\ f(\frac{1}{2}) = \frac{1}{2}$ and $f(x) = 1,\ x > \frac{1}{2}$ and is not in X. Consider a Cauchy sequences g_n and h_n in the equivalence class $\{f_n\}$;

$$g_n(x) = \begin{cases} 1 + n\,(x - \frac{1}{2}) & \text{on } [\frac{1}{2} - \frac{1}{n}, \frac{1}{2}] \\ 0 & \text{on } [0, \frac{1}{2} - \frac{1}{n}] \\ 1 & \text{on } [\frac{1}{2}, 1] \end{cases}$$

$$h_n(x) = \begin{cases} n\,(x - \frac{1}{2}) & \text{on } [\frac{1}{2}, \frac{1}{2} + \frac{1}{n}] \\ 0 & \text{on } [0, \frac{1}{2}] \\ 1 & \text{on } [\frac{1}{2} + \frac{1}{n}, 1]. \end{cases}$$

Thus, the candidate limits g, $(g(x) = 0,\ x < \frac{1}{2},\ g(x) = 1,\ x \ge \frac{1}{2}$ and h, $(h(x) = 0,\ x \le \frac{1}{2},\ h(x) = 1,\ x > \frac{1}{2}$ belong to the same equivalent class $f = \{f_n\}$. Note that f, g and h differ only at $x = \frac{1}{2}$. In general functions in an equivalent class differ at countably many points. The completion Y of X_2 is denoted by

$$L^2(0,1) = \{\text{the space of square integrable measurable functions on } (0,1)\}$$

with

$$|f|^2_{L^2} = \int_0^1 |f(x)|^2\,dx.$$

Here, the integrable is defined in the sense of Lebesgue as discussed in the next section.

The concept of completion states, for every $f \in Y = \bar{X}$ =the completion of X there exists a sequence f_n in X such that $|f_n - f|_Y \to 0$ as $n \to \infty$ and $|f|_Y = \lim_{n\to\infty} |f_n|_X$. For the example X_2, for all square integrable function $f \in L^2(0,1) = Y = \bar{X}_2$ there exists a continuous function sequence $f_n \in X_2$ such that $|f - f_n|_{L^2(0,1)} \to 0$.

Example $H^1_0(0,1)$ The completion of X_3 as defined above is denoted by

$$H^1_0(0,1) = \{\text{a space of absolutely continuous functions on } [0,1] \text{ with square integrable derivative and vanishing at } x = 0,\ x = 1\}$$

with

$$(f,g)_{H_0^1} = \int_0^1 \frac{d}{dx}f(x)\frac{d}{dx}g(x)\,dx$$

where the integrable is defined in the sense of Lebesgue. Here, f is absolutely continuous function if there exits an integrable function g such that

$$f(x) = f(0) + \int_0^x g(x)\,dx \quad x \in [0,1].$$

and f is almost everywhere (a.e.) differentiable and $\frac{d}{dx}f = g$ a.e. in $(0,1)$. For the linear spline function $B_k^n(x),\ 1 \le k \le n-1$ is in $H_0^1(0,1)$.

2.2 Measure Theory

In this section we discuss the measure theory.

Definition (1) A topological space is a set E together with a collection τ of subsets of E, called open sets and satisfying the following axioms; The empty set and E itself are open. Any union of open sets is open. The intersection of any finite number of open sets is open. The collection τ of open sets is then also called a topology on E. We assume that a topological space (E,τ) satisfies the Hausdoroff's axiom of separation:

for every disjoints $x_1,\ x_2 \in E$
there exist disjoint open sets $G_,\ G_2$ auch that $x_1 \in G_1,\ x_2 \in G_2$.

(2) A metric space E is a topological space with metric d if for all $x,\ y,\ z \in E$

$$d(x,y) \ge \text{ for } x,\ y \in E \text{ and } d(x,y) = 0 \text{ if and only if } x = y$$

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

$$d(x,z) \le d(x,y) + d(y,z).$$

For $x_0 \in E$ and $r > 0$ $B(x_0,r) = \{x \in E : d(x,x_0) < r$ is an open ball of E with center x_0 an radius r. A set $\mathcal{O}$ is called open if for any points $x \in E$ there exists an open ball $B(x,r)$ contained in the set $\mathcal{O}$.

(3) A collection of subsets $\mathcal{F}$ of E is σ-algebra if

$$\text{for all } A \in \mathcal{F},\quad A^c = E \setminus A \in \mathcal{F}$$

$$\text{for arbitrary family } \{A_i\} \text{ in } \mathcal{F}, \text{ countable union } \bigcup_{i=1}^\infty A_i \in \mathcal{F}$$

(4) A measure μ on $(E,\mathcal{F})$ assigns $A \in \mathcal{F}$ the nonnegative real value $\mu(A)$ and satisfies σ-additivity;

$$\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$$

for all disjoint sets $\{A_i\}$ in $\mathcal{F}$.

Theorem (Monotone Convergence) The measure μ is σ-additive if and only if for all sequence $\{A_k\}$ of nondecreasing events and $A = \bigcup_{k \ge 1} A_k$, $\lim_{n\to\infty} \mu(A_n) = \mu(A)$.

Proof: $\{A_k^c\}$ is a sequence of nonincreasing events and $A^c = \bigcap_{k\geq 1} A_k^c$. Since

$$\bigcap_{k\geq 1} A_k^c = A_1^c + (A_2^c \setminus A_1^c) + (A_3^c \setminus A_2^c) + \cdots$$

we have

$$\mu(A^c) = \mu(A_1^c) + \mu(A_2^c \setminus A_1^c) + \mu(A_3^c \setminus A_2^c) + \cdots$$

$$= \mu(A_1^c) + \mu(A_2^c) - \mu(A_1^c) + \mu(A_3^c) - \mu(A_1^c) + \cdots = \lim_{n\to\infty} \mu(A_n^c)$$

Thus,

$$\mu(A) = 1 - \mu(A^c) = 1 - (1 - \lim_{n\to\infty} \mu(A_n)) = \lim_{n\to\infty} \mu(A_n).$$

Conversely, let $A_1, A_2, \cdots \in \mathcal{F}$ be pairwise disjoint and let $\sum_{k=1}^\infty A_k \in \mathcal{F}$. Then

$$\mu(\sum_{k=1}^\infty A_k) = \mu(\sum_{k=1}^n A_k) + \mu(\sum_{k=n+1}^\infty A_k)$$

Since $\sum_{k=n+1}^\infty A_k \downarrow \varnothing$, we have

$$\lim_{n\to\infty} \sum_{k=n+1}^\infty A_k = \varnothing$$

and thus

$$\mu(\sum_{k=1}^\infty A_k) = \lim_{n\to\infty} \sum_{k=1}^n \mu(A_k) = \sum_{k=1}^\infty \mu(A_k).\square$$

Example (Metric space) A normed space X is a metric space with $d(x,y) = |x-y|$. A point $x \in E$ is the accumulation point of a sequence $\{x_n\}$ if $d(x_n, x) \to 0$ as $n \to \infty$. The closure $\bar{A}$ of a set A is the collection of all accumulation points of A. The boundary of asset A is the collection points of x that any open ball at x contains both points in A and A^c. A set B in E is closed if $B = \bar{B}$, i.e., every sequence in $\{x_n\}$ in B has a accumulation x. The complement A^c of an open set A is closed.
Examples (σ-algebra) There are trivial σ-algebras;

$$\mathcal{F}_0 = \{\Omega, \varnothing\}, \quad \mathcal{F}^* = \text{all subsets of } \Omega.$$

Let A be a subset of Ω and σ-algebra generated by A is

$$\mathcal{F}_A = \{\Omega, \varnothing, A, A^c\}.$$

A finite set of subsets $A_1, A_2, \cdots, A_n$ of E which are pairwise disjoint and whose union is E. It is called a partition of Ω . It generates the σ-algebra: $\mathcal{A} = \{A = \bigcup_{j\in J} A_j\}$ where J runs over all subsets of $1, \cdots, n$. This σ-algebra has 2^n elements. Every finite σ-algebra is of this form. The smallest nonempty elements $\{A_1, \cdots, A_n\}$ of this algebra are called atoms.

Example (Countable measure) Let Ω has a countable decomposition $\{D_k\}$, i.e.,

$$\Omega = \sum D_k, \quad D_j \cap D_j = \varnothing, \;\; i \neq j.$$

Let $\mathcal{F} = \mathcal{F}^*$ and $P(D_k) = \alpha_k > 0$ and $\sum_k \alpha_k = 1$. For the Poisson measure on nonnegative integers;

$$D_k = \{k\}, \quad m(D_k) = e^{-\lambda}\frac{\lambda^k}{k!}$$

for $\lambda > 0$.

Definition For any family C of subsets of E, we can define the σ-algebra $\sigma(C)$ by the smallest σ-algebra $\mathcal{A}$ which contains C. The σ-algebra $\sigma(C)$ is the intersection of all σ-algebras which contains C. It is again a σ-algebra, i.e.,

$$\mathcal{A} = \bigcap_\alpha \mathcal{A}_\alpha$$

where $\mathcal{A}_\alpha$ are all σ-algebras that contain C.

The following construction of the measure space and the measure is essential for the triple $(E, \mathcal{F}, m)$ on uncountable space E.

If $(E, \mathcal{O})$ is a topological space, where $\mathcal{O}$ is the set of open sets in E, then the σ-algebra $\mathcal{B}(E)$ generated by $\mathcal{O}$ is called the Borel σ-algebra of the topological space E and $(E, \mathcal{B}(E))$ defines the measure space and a set B in $\mathcal{B}(E)$ is called a Borel set. For example $(R^n, \mathcal{B}(R^n))$ is the the measure space and $\mathcal{B}(R^n)$ is the σ-algebra generated by open balls in R^n.

Caratheodory Theorem Let $\mathcal{B} = \sigma(\mathcal{A})$, the smallest σ-algebra containing an algebra $\mathcal{A}$ of subsets of E. Let μ_0 is a σ additive measure of on $(E, \mathcal{A})$. Then there exist a unique measure μ on $(E, \mathcal{B})$ which is an extension of μ_0, i.e., $\mu(A) = \mu_0(A), \ \ A \in \mathcal{A}$.

Definition A map f from a measure space $(X, \mathcal{A})$ to an other measure space $(Y, \mathcal{B})$ is called measurable, if $f^{-1}(B) = \{x \in X : f(x) \in B\} \in \mathcal{A}$ for all $B \in \mathcal{B}$. If $f : (R^n, \mathcal{B}(R^n)) \to (R^m, \mathcal{B}(R^m))$ is measurable, we say f is a Borel function.

In general every continuous function $R^n \to R^m$ is a Borel function since the inverse image of open sets in R^n are open in R^m.

Example (Measure space) Let $\Omega = R$ and $\mathcal{B}(R)$ be the Borel σ-algebra. Note that

$$(a, b] = \bigcap_n (a, b + \frac{1}{n}), \quad [a, b] = \bigcap_n (a - \frac{1}{n}, b + \frac{1}{n}) \in \mathcal{B}(R).$$

Thus, $\mathcal{B}(R)$ coincides with the σ-algebra generated by the semi-closed intervals. Let $\mathcal{A}$ be the algebra of finite disjoint sum of semi-closed intervals $(a_i, b_i]$ and define P_0 by

$$\mu_0(\sum_{k=1}^n (a_k, b_k]) = \sum_{k=1}^n (F(b_k) - F(a_k))$$

where $x \in R \to F(x) \in R^+$ is nondecreasing, right continuous and the left limit exists everywhere. We have the measure μ on $(R, \mathcal{B}(R))$ by the Caratheodory Theorem if μ_0 is σ-additive on $\mathcal{A}$. on $(\Omega, \mathcal{F}) = (R, \mathcal{B}(R))$.

We now prove that μ_0 is σ-additive on $\mathcal{A}$. By the monotone convergence theorem it suffices to prove that

$$P_0(A_n) \downarrow 0, \quad A_n \downarrow \varnothing, \quad A_n \in \mathcal{A}.$$

Without loss of the generality one can assume that $A_n \subset [-N, N]$. Since F is the right continuous, for each A_n there exists a set $B_n \in \mathcal{A}$ such that $\overline{B_n} \subset A_n$ and

$$\mu_0(A_n) - \mu_0(B_n) \le \epsilon 2^{-n}$$

for all $\epsilon > 0$. The collection of sets $\{[-N, N]\backslash\overline{B_n}\}$ is an open covering of the compact set $[-N, N]$ since $\cap\overline{B_n} = \varnothing$. By the Heine-Borel theorem there exists a finite subcovering:

$$\bigcup_{n=1}^{n_0} [-N, N] \setminus \overline{B_n} = [-N, N].$$

and thus $\bigcap_{n=1}^{n_0} \overline{B_n} = 0$. Thus,

$$\mu_0(A_{n_0}) = \mu_0(A_{n_0} \setminus \textstyle\bigcap_{k=1}^{n_0} B_k) + \mu_0(\bigcap_{k=1}^{n_0} B_k) = \mu_0(A_{n_0} \setminus \bigcap_{k=1}^{n_0} B_k)$$

$$\mu_0(\bigcap_{k=1}^{n_0} (A_k \setminus B_k)) \le \sum_{k=1}^{n_0} \mu_0(A_k \setminus B_k) \le \epsilon.$$

Since $\epsilon > 0$ is arbitrary $\mu_0(A_n) \to 0$ as $n \to \infty$.

Next, we define the integral of a measurable function f on $(\Omega, \mathcal{F}, \mu)$.

Defintion (Elementary function) An elementary function f is defined by

$$f(x) = \sum_{k=1}^{n} x_k \, I_{A_k}(x)$$

where $\{A_k\}$ is a partition of E, i.e, $A_k \in \mathcal{F}$ are disjoint and $\cup A_k = E$. Then the integral of f is given by

$$\int_E f(x)\, d\mu(x) = \sum_{k=1}^{n} x_k \mu(A_k).$$

Theorem (Approximation) For every measurable function $f \ge 0$ on $(E, \mathcal{F})$ there exists a sequence of elementary functions $\{f_n\}$ such that $0 \le f_n(x) \le f(x)$ and $f_n(x) \uparrow f(x)$ for all $x \in E$.

Proof: For $n \ge 1$, define a sequence of elementary functions by

$$f_n(\omega) = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{k,n}(\omega) + n I_{f(x)>n}$$

where $I_{k,n}$ is the indicator function of the set $\{x : \frac{k-1}{2^n} < f(x) \le \frac{k}{2^n}\}$. It is easy to verify that $f_n(x)$ is monotonically nondecreasing and $f_n(x) \le f(x)$ and thus $f_n(x) \to f(x)$ for all $x \in E$. $\square$

Definition (Integral) For a nonnegative measurable function f on $(E, \mathcal{F})$ we define the integral by

$$\int_E f(x)\, d\mu(x) = \lim_{n\to\infty} \int_E f_n(x)\, d\mu(x)$$

where the limit exists since $\int_E f_n(x)\,d\mu(x)$ is an increasing number sequence.

Note that $f = f^+ - f^-$ with $f^+(x) = \max(0, f(x))$, $f^-(x) = \max(0, f(x))$. So, we can apply for Theorem and Definition for f^+ and f^-.

$$\int_E f(x)\,d\mu(x) = \int_E f^+(x)\,d\mu(x) - \int_E f^-(x)\,d\mu(x)$$

If $\int_E f^+(x)\,d\mu(x)$, $\int_E f^-(x)\,d\mu(x) < \infty$, f is integrable and

$$\int_E |f(x)|\,d\mu(x) = \int_E f^+(x)\,d\mu(x) + \int_E f^-(x)\,d\mu(x).$$

Corollary Let f is a measurable function on $(R, \mathcal{B}(R), \mu)$ with $\mu((a,b]) = F(b) - F(a)$, then we have as $n \to \infty$

$$\int_0^\infty f(x)\,d\mu(x) = \sum_{k=0}^{n2^n} f(\frac{k-1}{2^n})(F(\frac{k}{2^n}) - F(\frac{k-1}{2^n})) + f(n)(1 - F(n)) \to \int_0^\infty f(x)\,dF(x),$$

which is the Lebesgue Stieljes integral with respect to measure dF.

The space $\mathcal{L}$ of measurable functions on $(E, \mathcal{F})$ is a complete metric space with metric

$$d(f,g) = \int_E \frac{|f-g|}{1+|f-g|}\,d\mu.$$

In fact, $d(f,g) = 0$ if and if $f = g$ almost everywhere and since

$$\frac{|f+g|}{1+|f+g|} \le \frac{|f|}{1+|f|} + \frac{|g|}{1+|g|},$$

d satisfies the triangle inequality.

Theorem (completeness) $(\mathcal{L}, d)$ is a complete metric space.

Proof: Let $\{f_n\}$ be a Cauchy sequence of $(\mathcal{L}, d)$. Select a subsequence f_{n_k} such that

$$\sum_{k=1}^\infty d(f_{n_k}, f_{n_{k+1}}) < \infty$$

Then,

$$\sum_k E[\frac{|f_{n_k} - f_{n_{k+1}}|}{1+|f_{n_k} - f_{n_{k+1}}|} < \infty$$

Since $|f| \le\le \frac{2|f|}{1+|f|}$ for $|f| \le 1$,

$$\sum^\infty |f_{n_k} - f_{n_{k+1}}| < \infty \text{ a.e.}$$

and $\{f_{n_k}\}$ almost everywhere converges to f for some $f \in \mathcal{L}$. Moreover, $d(f_{n_k}, f) \to 0$ and thus $(\mathcal{L}, d)$ is complete.

Definition (Convergence in Measure) A sequence $\{f_n\}$ of mesurable functions on S converges to f in measure if

$$\text{for } \epsilon > 0,\ \lim_{n\to\infty} m(\{s \in S : |f_n(s) - f(s)| \geq \epsilon\}){=}0.$$

Theorem (Convergence in measure) $\{f_n\}$ converges in measure to f if and only if $\{f_n\}$ converges to f in d-metric.

Proof: For $f,\ g \in \mathcal{L}$

$$d(f,g) = \int_{|f-g|\geq\epsilon} \frac{|f-g|}{1+|f-g|}\,d\mu + \int_{|f-g|<\epsilon} \frac{|f-g|}{1+|f-g|}\,d\mu \leq \mu(|f-g| \geq \epsilon) + \frac{\epsilon}{1+\epsilon}$$

holds for all $\epsilon > 0$. Thus, f_n converges in mesure to f, then f_n converges to f in d-metric. Conversely, since

$$d(f,g) \geq \frac{\epsilon}{1+\epsilon}\mu(|f-g| \geq \epsilon)$$

if f_n converges to f in d-metric, then f_n converges in measure to f. $\square$

Corollary The completion of $C(E)$ with respect to $L^p(E)$-norm:

$$\left(\int_E |f|^p\,d\mu\right)^{\frac{1}{p}}$$

is $L^p(E) = \{f \in \mathcal{L} : \int_E |f|^p\,d\mu\}$.

2.3 Baire's Category Theory

In this section we discuss the uniform boundedness principle and the open mapping theorem. The underlying idea of the proofs of these theorems is the Baire theorem for complete metric spaces. It is concerned with the decomposition of a space as a union of subsets. For instance, we have a countable union;

$$R^2 = \bigcup_{k,j} S_{k,j}, \quad S_k = (k, k+1] \times (j, j+1].$$

On the other hand, we have uncountable union;

$$\bigcup_{\alpha\in R} \ell_\alpha, \quad \ell_\alpha = \{\alpha\} \times R$$

but ℓ_α has no interior points, the one dimensional subspace of R^2. The question is can we represent R^2 as a countable union of these sets? It turns out that the answer is no.

Definition (Category) Let (X,d) be a metric space. A subset E of X is called nowhere dense if its closure does not contain any open set. X is said to be the first category if X is the union of a countable number of nowhere dense sets. Otherwise, X is said to be the second category, i.e., suppose $X = \bigcup_k E_k$, then the closure of at least one of E_k's has non-empty interior.

Baire-Hausdorff Theorem A complete metric space is of the second category.

Proof: Let $\{E_k\}$ be a sequence of closed sets and assume $\cup_n E_n = X$ in a complete metric space X. We will show that there is at least one of E_k's has non-empty interior. Assume otherwise, no E_n contains an interior point. Let $O_n = E_n^c$ and then O_n is open and dense in X for every $n \geq 1$. Thus, O_1 contains a closed ball $B_1 = \{x : d(x, x_1) \leq r_1\}$ with $r_1 \in (0, \frac{1}{2})$. Since O_2 is open and dense, there exists a closed ball $B_2 = \{x : d(x, x_2) \leq r_2\}$ with $r_2 \in (0, \frac{1}{2^2})$ in B_1. By repeating the same process, there exists a sequence $\{B_k\}$ of closed ball $B_k = \{x : x : d(x, x_k) \leq r_k$ such that

$$B_{k+1} \subset B_k, \;\; 0 < r_k \frac{1}{2^k}, \;\; B_k \cap E_k = \varnothing.$$

The sequence $\{x_k\}$ is a Cauchy sequence in X and since X is complete $x^* = \lim x_n \in X$ exists. Since

$$d(x_n, x^*) \leq d(x_n, x_m) + d(x_m, x^*) \leq r_n + d(x_m, x^*) \to r_n$$

as $m \to \infty$, $x^{\in} B_n$ for every n, i.e. $x^* \notin E_n$ and thus $x^* \notin \cap_n E_n = X$, which is a contradiction. □

Baire Category theorem Let M be a set of the first category in a compact topological space. Then, the complement M^c is dense in X.

Theorem (Totally bounded) A subset M in a complete metric space X is relatively compact if and only if it is totally bounded, i.e, for every $\epsilon > 0$ there exists a family of finite many points $\{m_1, m_2, \cdots, m_n\}$ in M such that for every point $m \in M$, $d(m, m_k) < \epsilon$ for at least one m_k.

Theorem (BanachSteinhaus, uniform boundedness principle) Let $X,\ Y$ be Banach spaces and let $T_k,\ k \in K$ be a family (not necessarily countable) of continuous linear operators from X to Y. Assume that

$$\sup_{k \in K} |T_k x|_Y < \infty \text{ for all } x \in X.$$

Then there exists a constant c such that

$$|T_k x|_Y \leq c\,|x|_X \text{ for all } k \in K \text{ and } x \in X.$$

Proof: Define

$$X_n = \{x \in X : |T_k x| \leq n \text{ for all } k \in K\}.$$

The, X_n is closed, and by the assumption we have $X = \cup_n X_n$. It follows from the Baire category theorem that $int(X_0)$ is not empty for some n_0 Pick $x_0 \in X$ and $r > 0$ such that $B(x_0, r) \subset int(X_0)$. We have

$$T_k(x_0 + r\,z) \leq n_0 \text{ for all } k \text{ and } z \in B(0,1),$$

which implies

$$r\,|T_k|_{\mathcal{L}(X,Y)} \leq n_0 + |T_k x_0|.\square$$

The uniform bounded principle is quite remarkable and surprising, since from pointwise estimates one derives a global (uniform) estimate. We have the following examples.

Examples (1) Let X be a Banach space and B be a subset of X. If for every $f \in X^*$ the set $f(B) = \{\langle f, x\rangle,\ x \in B\}$ is bounded in R, then B is bounded.

(2) Let X be a Banach space and B^* be a subset of X^*. If for every $x \in X$ the set $\langle B^*, x\rangle = \{\langle f, x\rangle,\ f \in B^*\}$ is bounded in R, then B^* is bounded.

Theorem (Banach closed graph theorem) A closed linear operator on a Banach space X to a Banach space Y is continuous.

2.4 Dual space and Hahn-Banach theorem

In this section we discuss the dual space of a normed space and the Hahn-Banach theorem.

Definition (Dual Space) If X be a normed space, let X^* denote the collection of all bounded linear functionals on X. X^* is called the dual space of X. We define the dual product $\langle x, f\rangle_{X\times X^*}$ by

$$\langle f, x\rangle_{X^*\times X} = f(x) \quad \text{for } \ x \in X \ \ f \in X^*.$$

For $f \in X^*$ we define the operator norm of f by

$$|f|_{X^*} = \sup_{x\neq 0} \frac{|f(x)|}{|x|_X}.$$

Theorem (Dual space) The dual space X^* equipped with the operator norm is a Banach space.

Proof: $\{f_n\}$ be a Cauchy sequence in X^*, i.e.

$$|f_n(\phi) - f_m(\phi)| \leq |f_n - f_m|_*|\phi| \to 0 \text{ as } m \to \infty,$$

for all $\phi \in X$. Thus, $\{f_n(\phi)\}$ is Cauchy in R, define the functional f on X by

$$f(\phi) = \lim_{n\to\infty} f_n(\phi), \quad \phi \in X.$$

Then f is linear and for all $\epsilon > 0$ there exists N_ϵ such that for $m,\ n \geq N_\epsilon$

$$|f_m(\phi) - f_n(\phi)| \leq \epsilon\,|\phi|$$

Letting $m \to \infty$, we have

$$|f(\phi) - f_n(\phi)| \leq \epsilon\,|\phi| \text{ for all } n \geq N_\epsilon.$$

and thus $f \in X^*$ and $|f_n - f|_* \to 0$ as $n \to \infty$. $\square$

Examples (Dual space) (1) Let ℓ^p be the space of p-summable infinite sequences $x = (s_1, s_2, \cdots)$ with norm $|x|_p = (\sum |s_k|^p)^{\frac{1}{p}}$. Then, $(\ell^p)^* = \ell^q$, $\frac{1}{p} + \frac{1}{q} = 1$ for $1 \leq p < \infty$. Let $c = \ell^\infty$ be the space of uniformly bounded sequences $x = (\xi_1, \xi_2, \cdots)$ with norm $|x|_\infty = \max_k |\xi_k|$ and c_0 be the space of uniformly bounded sequences with $\lim n \to \infty \xi_n = 0$. Then, $(\ell^1)' = c_0$
(2) Let $L^p(\Omega)$ is p-integrable functions f on Ω with

$$|f|_{L^p} = (\int_\Omega |f(x)|^p\,dx)^{\frac{1}{p}}$$

Then, $L^p(\Omega)^* = L^q(\Omega)$, $\frac{1}{p}+\frac{1}{q}=1$ for $p \geq 1$. Especially, $L^2(\Omega)^* = L^2(\Omega)$ and $L^1(\Omega)^* = L^\infty(\Omega)$, where

$$L^\infty(\Omega) = \text{ the space of essentially bounded functions on } \Omega$$

with

$$|f|_{L^\infty} = \inf M \text{ where } |f(x)| \leq M \text{ almost everywhere on } \Omega.$$

(3) Consider a linear functional $f(x) = x(1)$. Then, f is in X_1^* but not in X_2^* since for the sequence defined by $x^n(t) = t^n$, $t \in [0,1]$ we have $f(x_n) = 1$ but $|x^n|_{X^2} = \sqrt{\frac{1}{2n+1}} \to 0$ as $n \to \infty$. Since for $x \in X_3$

$$|x(t)| = |x(0) + \int_0^t x'(s)\,ds| \leq \sqrt{\int_0^1 |x'(t)|^2\,dt},$$

thus $f \in X_3^*$.

(4) The total variation $V_0^1(g)$ of a function g on an interval $[0,1]$ is defined by

$$V_0^1(g) = \sup_{P\in\mathcal{P}} \sum_{i=0}^{n-1} |g(x_{i+1}) - g(x_i)|,$$

where the supremum is taken over all partitions $\mathcal{P} = \{P : 0 \leq t_0 \leq \cdots \leq t_n \leq 1\}$ of the interval $[0,1]$. Then, $BV(0,1) = \{$uniformly bounded function g with $V_0^1(f) < \infty\}$ and $|g|_{BV} = V_0^1(g)$. It is a normed space if we assume $g(0) = 0$. For example, every monotone functions is of bounded variation. Every real-valued BV-function can be expressed as the difference of two increasing functions (Jordan decomposition theorem).

<u>Hahn-Banach Theorem</u> Suppose S is a subspace of a normed space X and $f : S \to R$ is a linear functional satisfying $f(x) \leq |x|$ for all $x \in S$. Then there exists an $F \in X^*$ such that

$$F(x) \leq |x| \quad \text{and} \quad F(s) = f(s) \text{ for all } s \in S.$$

In general, the theorem holds for a vector space X and sub-additive and positive homogeneous function p, i.e., for all $x,\ y \in X$ and $\alpha > 0$

$$p(x+y) \leq p(x) + p(y), \quad p(\alpha x) = \alpha p(x).$$

Every linear functional f on a subspace S, satisfying $f(x) \leq p(x)$ has a extension F on X satisfying $F(x) \leq p(x)$ for all $x \in X$.

Proof: First, we show that there exists an one step extension of f, i.e., for $x_0 \in X \setminus S$, there exists an extension f_1 of f on the subspace Y_1 spanned by (S, x_0) such that $f_1(x) \leq p(x)$ on Y_1. Note that for arbitrary $s_1,\ s_2 \in S$,

$$f(s_1) + f(s_2) \leq p(s_1 - x_0) + p(s_2 + x_0).$$

Hence, there exist a constant c such that

$$\sup_{s\in S}\{f(s) - p(s - x_0)\} \leq c \leq \inf_{s\in S}\{p(s + x_0) - f(s)\}.$$

Extend f by defining, for $s \in S$ and $\alpha \in R$

$$f_1(s + \alpha x_0) = f(s) + \alpha\, c,$$

where $f_1(x_0) = c$, is determined as above. We claim that

$$f_1(s + \alpha x0) \le p(s + \alpha x_0)$$

for all $\alpha \in R$ and $s \in S$, i.e., by the definition of c, for $\alpha > 0$

$$f_1(s + \alpha x_0) = \alpha(c + f(\frac{s}{\alpha}) \le \alpha\,(p(\frac{s}{\alpha} + x_0) - f(\frac{s}{\alpha}) + f(\frac{s}{\alpha}) = p(s + \alpha x_0),$$

and for all $\alpha < 0$

$$f_1(s + \alpha x_0) = -\alpha(-c + f(\frac{s}{-\alpha}) \le -\alpha(p(\frac{s}{-\alpha} - x_0) - f(\frac{s}{-\alpha}) + f(\frac{s}{-\alpha}) = p(s + \alpha x_0).$$

If X is a separable normed space. Let $\{x_1, x_2, \dots\}$ be a countable dense base of $X \setminus S$. Select victors from this dense vectors, which is independent of S, recursively. Extend f to subspaces $span\{span\{span\{S, x_1\}, x_2\}, \dots\}$ recursively, as described above. The space $\mathrm{span}\{S, x_1, x_2, \dots\}$ is dense in X. The final extension F is defined by the continuity. In general, the remaining proof is done by using Zorns lemma. $\square$

Corollary Let X be a normed space. For each $x_0 \in X$ there exists an $f \in X^*$ such that

$$f(x_0) = |f|_{X^*}|x_0|_X.$$

Proof of Corollary: Let $S = \{\alpha\, x_0 : \alpha \in R\}$ and define $f(\alpha\, x_0) = \alpha\, |x_0|_X$. By Hahn-Banach theorem there exits an extension $F \in X^*$ of f such that $F(x) \le |x|$ for all $x \in X$. Since

$$-F(x) = F(-x) \le |-x| = |x|,$$

we have $|F(x)| \le |x|$, in particular $|F|_{X^*} \le 1$. On the other hand, $F(x_0) = f(x_0) = |x_0|$, thus $|F|_{X^*} = 1$ and $F(x_0) = f(x_0) = |F||x_0|$. $\square$

Example (Hahn-Banach Theorem) (1) $|x| = \sup_{|x^*|=1} \langle x^*, x\rangle$
(2)

Next, we prove the geometric form of the Hahn-Banach theorem. It states that given a convex set K containing an interior point and $x_0 \notin K$, then x_0 is separated from K by a hyperplane. Let K be a convex set and 0 is an interior point of K. A hyperplane is a maximal affine space $x_0 + M$ where M is a subspace of X. We introduce the Minkowski functional that defines the distance from the origin measured by K.

Definition (Minkowski functional) Let K be a convex set in a normed space X and 0 is an interior point of K. Then the Minkowski functional p of K is defined by

$$p(x) = \inf\{r : \frac{x}{r} \in K,\ r > 0\}.$$

Note that if K is the unit sphere in X, $p(x) = |x|$.

Lemma (Minkowski functional)
(1) $0 \le p(x) < \infty$.
(2) $p(\alpha x) = \alpha\, p(x)$ for $\alpha \ge 0$.

(3) $p(x_1 + x_2) \leq p(x_1) + p(x_2)$.
(4) p is continuous.
(5) $\bar{K} = \{x : p(x) \leq 1\}$, $int(K) = \{x : p(x) < 1\}$.

Proof: (1) Since K contains a sphere at 0, for every $x \in X$ there exists an $r > 0$ such that $\frac{x}{r} \in K$ and thus $p(x)$ is finite.
(2) For $\alpha > 0$

$$p(\alpha x) = \inf\{r : \frac{\alpha x}{r} \in K\} = \inf\{\alpha r' : \frac{x}{r'} \in K\} = \alpha \inf\{r' : \frac{x}{r'} \in K\} = \alpha\, p(x).$$

(3) Let $p(x_i) < r_i < p(x_i)+\epsilon,\ i = 1,\ 2$ for arbitrary $\epsilon > 0$. Let $r = r_1+r_2$. By convexity of K since $\frac{r_1}{r}\frac{x_1}{r_1} + \frac{r_2}{r}\frac{x_2}{r_2} = \frac{x_1+x_2}{r} \in K$ and thus $p(x_1 + x_2) \leq r \leq p(x_1) + p(x_2) + 2\epsilon$. Since $\epsilon > 0$ is arbitrary, p is subadditive.
(4) Let ϵ be a radius of a closed sphere centered at 0 and contained in K. Since $\epsilon\frac{x}{|x|} \in K$, $p(\epsilon\frac{x}{|x|}) \leq 1$ and thus $p(x) \leq \frac{1}{\epsilon}|x|$. This shows that p is continuous at 0. Since p is sub linear

$$p(x) = p(x - y + y) \leq p(x - y) + p(y) \text{ and } p(y) = p(y - x + x) \leq p(y - x) + p(x)$$

and thus

$$-p(y - x) \leq p(x) - p(y) \leq p(x - y).$$

Hence, p is continuous since it is continuous at 0.
(5) It follows from the continuity of p. $\Box$

Lemma (Hyperplane) Let H be a hyperplane in a vector space X if and only if there exist a linear functional f on X and a constant c such that $H = \{f(x) = c\}$.

Proof: If H ia hyperplane, then H is the translation of a subspace M in X, i.e., $H = x_0 + M$. If $x_0 \not inM$, then $[M + x_0] = X$. if for $x = \alpha\, x_0 + m,\ m \in M$ we let $f(x) = \alpha$, we have $H = \{x : f(x) = 1\}$. Conversely, if f is a nonzero linear functional on X, then $M = \{x : f(x) = 0\}$ is a subspace. Let $x_0 \in X$ such that $f(x_0) = 1$. Since $f(x - f(x)x_0) = 0$, $X = span[\{x_0\}, M]$ and $\{x : f(x) = c\} = \{f(x - c\, x_0) = 0\}$ is a hyperplane.
Theorem (Mazur's Theorem) Let K be convex set with a nonempty interior point in a normed space X. Suppose V is an affine space in X that contains no interior points of K. Then, there exist $x^* \in X^*$ and constant c such that the hyperplane $langlex^*, v\rangle = c$ has the separation:

$$\langle x^*, v\rangle = c \text{ for all } v \in V \text{ and } \langle x^*, k\rangle < c \text{ for all } k \in int(K).$$

Proof: By an appropriate translation one may assume that 0 is an interior point of K. Let M is the subspace spanned by V. Since V does not contain 0, exists a linear functional f on M such that $V = \{x : f(x) = 1\}$. Let p be the Minkowski functional of K. Since V contains no interior points of K, $f(x) = 1 \leq p(x)$ for all $x \in V$. By the homogeneity of p, $f(x) \leq p(x)$ for all $x \in V$ By the Hahn-Banach space, there exits an extension F of f to V from M to X with $F(x) \leq p(x)$. From Lemma p is continuous and F is continuous and $F(x) < 1$ for $x \in int(K)$. Thus, $H = \{x : F(x) = 1\}$ is the desired closed hyperplane. $\Box$

Example (Hilbert space) If X is a Hilbert space and S is a subspace of X. Then the extension F is defined as

$$F(s) = \begin{cases} \lim f(s_n) & \text{for } s \in \overline{S} \\ 0 & \text{for } s \in S^\perp \end{cases}$$

where $X = \overline{S} \oplus S^\perp$.
Example (Integral) Let $X = L^2(0,1)$ $S = C(0,1)$ and $f(x) = R - \int_0^1 x(t)\,dt$ be the Riemman integral of $x \in S$. Then the extension F is $F(x) = L - \int_0^1 x(t)\,dt$ is the Lebesgue integral of $x \in X$.

Example (Alignment) (1) For Hilbert space X $f = x_0$ is self aligned.
(2) Let $x_0 \in X = L^p(\Omega),\ 1 \le p < \infty$. For $f \in X^* = L^q(\Omega)$ defined by by

$$f(t) = |x_0(t)|^{p-2}x_0(t), \quad t \in \Omega,$$

$$f(x) = \int_\Omega |x_0(t)|^p\,dt.$$

Thus, $\frac{x_0}{|x_0|^{p-2}} \in X^*$ is aligned with x_0.

A function g is called right continuous if $\lim_{h\to 0^+} g(t+h) = g(t)$. Let

$$BV_0[0,1] = \{g \in BV_0[0,1] : g \text{ is right continuous on } [0,1) \text{ and } g(0) = 0\}.$$

Let $C[0,1]$ is the normed space of continuous functions on $[0,1]$ with the sup norm. For $f \in C[0,1]$ and $g \in BV_0[0,1]$ define the Riemann-Stieltjes integral by

$$\int_0^t x(t)dg(t) = \lim_{\Delta t} R(f,g,P) = \lim_{\Delta t} \sum_{k=1} f(z_k)(g(t_k) - g(t_{k-1})$$

for the partition $P = \{0 = t_0 < \cdots < t_n = 1\}$ and $z_k \in [t_{k-1}, t_k]$. There is a version of the Riesz representation theorem.

Theorem (Dual of $C[0,1]$) There is a norm-preserving linear isomorphism from $C[0,1]$ to $V[0,1]$, i.e., for $F \in C[0,1]^*$ there exists a unique $g \in BV_0[0,1]$ such that

$$F(x) = \int_0^1 x(t)dg(t)$$

and

$$|F|_{C[0,1]^*} = |g|_{BV}.$$

That is, $C[0,1]^* \cong BV(0,1)$, the space of functions with bounded variation.

Proof: Define the normed space of uniformly bound function on $[0,1]$ by

$$B = \{x : [0,1] \to R \text{ is bounded on } [0,1]\} \text{ with norm } |x|_B = \sup_{t\in[0,1]} |x(t)|.$$

By the Hahn-Banach theorem, an extension F of f from $C[0,1]$ to B exists and $|F| = |f|$. For any $s \in (0,1]$, define $\chi_{[0,s]} \in B$ and define $v(s) = F(\chi_{[0,s]})$. Then,

$$\sum_k |v(t_k) - v(t_{k-1})| = \sum_k sign(v(t_k) - v(t_{k-1}))(v(t_k) - v(t_{k-1}))$$

$$= sign(v(t_k) - v(t_{k-1})(F(\chi_{[0,t_k]}) - F(\chi_{[0,t_{k-1}]})) = F(\sum_k sign(v(t_k) - v(t_{k-1})(\chi_{[0,t_k]} - \chi_{[0,t_{k-1}]})$$

$$\le |F||\sum_k sign(v(t_k) - v(t_{k-1})\chi_{(t_{k-1},t_k]}| \le |f|.$$

Thus, $v \in BV[0,1]$ and $|v|_{BV} \leq |F|$. Next, for any $x \in C[0,1]$ define the function

$$z(t) = \sum_k x(t_{k-1}(\chi_{[0,t_k]} - \chi_{[0,t_{k-1}]}).$$

Note that the difference

$$|z - x|_B = \max_k \max_{t \in [t_{k-1},t_k]} |x(t_{k-1}) - x(t)| \to 0 \text{ as } \Delta t \to 0^+.$$

Since F is bounded $F(z) \to F(x) = f(x)$ and

$$F(z) = \sum_k x(t_{k-1})(v(t_k) - v(t_{k-1})) \to \int_0^1 x(t)dv(t)$$

we have

$$f(x) = \int_0^1 x(t)dv(t)$$

and $|F| \leq |x|_{C[0,1]}|v|_{BV}$. Hence $F| = |f| = |v|_{BV}$. □

Riesz Representation Theorem For every bounded linear functional f on a Hilbert space X there exits a unique element $y_f \in X$ such that

$$f(x) = (x, y_f) \text{ for all } x \in X, \text{ and } |f|_{X^*} = |y_f|_X.$$

We define the Riesz map $R : X^* \to X$ by

$$Rf = y_f \in X \quad \text{for } f \in X^*$$

for a Hilbert space X.
Proof: There exists a $z \in X$ such that $(z, x) = 0$ for all $f(x) = 0$ and $|z| = 1$. Otherwise, $f(x) = 0$ for all x and $f = 0$. Note that for every $x \in X$ we have $f(f(x)z - f(z)x) = 0$. Thus,

$$0 = (f(x) - f(z)x, z) = (f(x)z, z) - (f(z)x, z)$$

which implies

$$f(x) = (x, \overline{f(z)}z)$$

The existence then follows by taking $y_f = \overline{f(z)}z$. Let for $u_1,\ u_2$ $f(x) = (x, u_1) = (x, u_2)$ for all $x \in X$. Then, then $(x, u_1 - u_2) = 0$ and by taking $x = u_1 - u_2$ we obtain $|u_1 - u_2| = 0$, which implies the uniqueness.

Definition (Reflexive Banach space) For a Banach space X we define an injection i of X into $X^{**} = (X^*)^*$ (the double dual) as follows. If $x \in X$ then $i\,x \in X^{**}$ is defined by $i\,x(f) = f(x)$ for $f \in X^*$. Note that

$$|i\,x|_{X^{**}} = \sup_{f \in X^*} \frac{|f(x)|}{|f|_{X^*}} \leq \sup_{f \in X^*} \frac{|f|_{X^*}|x|_X}{|f|_{X^*}} = |x|_X$$

an thus $|i\,x|_{X^**} \leq |x|_X$. But by Hahn-Banach theorem there exists an $f \in X^*$ such that $|f|_{X^*} = 1$ and $f(x) = |x|_X$. Hence $|i\,x|_{X^**} = |x|_X$ and $i : X \to X^{**}$ is an isometry. If i is also subjective, then we say that X is reflexive and we may identify X with X^{**}.

Definition (Weak and Weak star Convergence) (1) A sequence $\{x_n\}$ in a normed space X is said to converge weakly to $x \in X$ if $\lim_{n\to\infty} f(x_n) \to f(x)$ for all $f \in X^*$
(2) A sequence $\{f_n\}$ in the dual space X^* of a normed space X is said to converge weakly star to $f \in X^*$ if $\lim_{n\to\infty} f_n(x) \to f(x)$ for all $x \in X$.

Notes (1) If x_n converges strongly to $x \in X$ then x_n converges weakly to x, but not conversely.
(2) If x_n converges weakly to x, then $\{|x_n|_X\}$ is bounded and $|x| \leq \liminf |x_n|_X$, i.e., the norm is weakly lower semi-continuous.
(3) A sequence $\{x_n\}$ in a normed space X converges weakly to $x \in X$ if $\lim_{n\to\infty} f(x_n) \to f(x)$ for all $f\ inS$, where S is a dense set in X^*.

(4) If X be a Hilbert space. If a sequence $\{x_n\}$ of X converges weakly to $x \in X$, then x_n converges strongly to x if and only if $\lim_{n\to\infty} |x_n| = |x|$. In fact,

$$|x_n - x|^2 = (x_n - x, x_n - x) = |x_n|^2 - (x_n, x) - (x, x_n) + |x|^2 \to 0$$

as $n \to \infty$ if x_n converges weakly to x.

Example (Weak Convergence) (1) Let $X = L^2(0,1)$ and consider a sequence $x_n = \sqrt{2}\,\sin(n\,\pi t),\ t \in (0,1)$. Then $|x_n| = 1$ and x_n converges weakly to 0. In fact, since a family $\{\sin(k\pi t)\}$ is dense in $L^2(0,1)$ and

$$(x_n, \sin(k\pi t))_{L^2} \to 0 \text{ as } n \to \infty$$

for all k and thus x_n converges weakly to 0. Since $|0| = 0$, the sequence x_n does not converge strongly to 0.

In general, if $x_n \to x$ weakly and $y_n \to y$ strongly in a Hilbert space, then $(x_n, y_n) \to (x, y)$ since

$$(x_n, y_n) = (x_n - x, y) + (x_n, y_n - y).$$

But, note that $(x_n, x_n) = 2\int_0^1 \sin^2(n\pi t)\,dt = 1 \neq 0$ and (x_n, x_n) does not converges to 0. In general (x_n, y_n) does not necessarily converge to (x, y) if $x_n \to x$ and $y_n \to y$ weakly in X.

Alaoglu Theorem If X is a normed space and $\{x_n^*\}$ is a bounded sequence in X^*, then there exists a subsequence that converges weakly star to an element of X^*.

Proof: We prove the theorem for the case when X is separable. Let $\{x_n^*\}$ be sequence in X^* such that $|x_n^*| \leq 1$. Let $\{x_k\}$ is a dense sequence in X. The sequence $\{\langle x_n^*, x_1\rangle\}$ is a bounded sequence in R and thus a convergent subsequence denoted by $\{\langle x_{n_1}^*, x_1\rangle\}$. Similarly, the sequence $\{\langle x_{n_1}^*, x_2\rangle\}$ contains a a convergent subsequence $\{\langle x_{n_2}^*, x_1\rangle\}$. Continuing in this manner we extract subsequences $\{\langle x_{n_k}^*, x_k\rangle\}$ that converges in R. Thus, the diagonal sequence $\{x_{nn}^*\}$ on the dense subset $\{x_k\}$, i.e., $\{\langle x_{nn}^*, x_k\rangle\}$ converges for each x_k. We will show that x_{nn}^* converges weakly start to an element $x^* \in X^*$. Let a $x \in X$ be fixed. For arbitrary $\epsilon > 0$ there exists $N \geq N_\epsilon$ such that

$$|x - \sum_{k=1}^N \alpha_k\, x_k| \leq \frac{\epsilon}{3}.$$

Thus,

$$|\langle x_{nn}^*, x\rangle - \langle x_{mm}^*, x\rangle| \leq |\langle x_{nn}^*, x - \textstyle\sum_{k=1}^N \alpha_k x_k\rangle|$$

$$+|\langle x_{nn}^* - x_{mm}^*, \textstyle\sum_{k=1}^N \alpha_k x_k\rangle| + |\langle x_{mm}^*, x - \textstyle\sum_{k=1}^N \alpha_k x_k\rangle| \leq \epsilon$$

for sufficiently large $nn,\ mm$. Thus, $\{\langle x^*_{nn}, x\rangle\}$ is a Cauchy sequence in R and converges to a real number $\langle x^*, x\rangle$. The functional $\langle x^*, x\rangle$ so defined is obviously linear and $|x^*| \le 1$ since $|\langle x^*_{nn}, x\rangle| \le |x|$. $\square$

Example (Weak Star Convergence) (1) Let $X = L^1(\Omega)$ and $L^\infty(\Omega) = X^*$. Thus, a bounded sequence in $L^\infty(\Omega)$ has a weakly star convergent subsequence. Let $X = L^\infty(0,1)$ and $\{x_n\}$ be a sequence in X defined by

$$x_n(t) = \begin{cases} 1 & \text{on } [\frac{1}{2}+\frac{1}{n}, 1] \\ \\ n\,x & \text{on } [\frac{1}{2}-\frac{1}{n}, \frac{1}{2}-\frac{1}{n}] \\ \\ -1 & \text{on } [0, \frac{1}{2}-\frac{1}{n}]; \end{cases}$$

Then $|x_n|_X = 1$ and $\{x_n\}$ converges weakly star to the sign function $x(t) = sign(t-\frac{1}{2})$. But, since $|x_n - x| = \frac{1}{2}$, $\{x_n\}$ does not converge strongly to x.
(2) Let $X = c_0$, the space of infinite sequences $x = (x_1, x_2, \cdots)$ that converge to zero with norm $|x|_\infty = \max_k |x_k|$. Then $X^* = \ell_1$ and $X^{**} = \ell_\infty$. Let $\{x^*_n\}$ in $\ell_1 = X^*$ defined by $x^*_n = (0, \cdots, 1, 0, \cdots)$ where 1 appears at $n-$th term only. Then, x^*_n converge weakly star to 0 in X^* but x^*_n does not converge weakly to 0 since $\langle x^*_n, x^{**}\rangle$ does not converge for $x^{**} = (1, 1, \cdots) \in X^{**}$.

Example ($L^1(\Omega)$ bounded sequence) Let $X = L^1(0,1)$ and let $\{x_n\} \in X$ be a sequence defined by

$$x_n(t) = \begin{cases} n & t \in (\frac{1}{2}-\frac{1}{2n}, \frac{1}{2}+\frac{1}{2n}) \\ 0 & \text{otherwise.} \end{cases}$$

Then, $|x_n|_{L^1} = 1$ and

$$\int_0^1 x_n(t)\phi(t)\,dt \to \phi(\frac{1}{2})$$

for all $\phi \in C[0,1]$ but a linear functional $f(\phi) = \phi(\frac{1}{2})$ is not bounded on $L^\infty(0,1)$. Thus, x_n does not converge weakly. Actually, sequence x_n converges weakly star in $C[0,1]^*$.

Eberlein-Suhmulyan Theorem A Banach space X is reflexive if and only if and only if every closed bounded set is weakly sequentially compact, i.e., every bounded sequence of X contains a subsequence that converges weakly to an element of X.

Next, we consider linear operator $T : X \to Y$.

Definition (Linear Operator) Let X, Y be normed spaces. A linear operator T on $\mathcal{D}(T) \subset X$ into Y satisfies

$$T(\alpha\, x_1 + \beta\, x_2) = \alpha\, Tx_1 + \beta\, Tx_2$$

where the domain $\mathcal{D}(T)$ is a subspace of X. Then a linear operator T is continuous if and only if there exists a constant $M \ge 0$ such that

$$|Tx|_Y \le M\,|x|_X \quad \text{for all } x \in X.$$

For a continuous linear operator we define the operator norm $|T|$ by

$$|T| = \sup_{|x|_X = 1} |Tx|_Y = \inf M,$$

where we assume $\mathcal{D}(T) = X$. A continuous linear operator on a normed space X into Y is called a *bounded linear operator* on X into Y and will be denoted by $T \in \mathcal{L}(X, Y)$. The norm of $T \in \mathcal{L}(X, Y)$ is denoted either by $|T|_{\mathcal{L}(X,Y)}$ or simply by $|T|$. Note that the space $\mathcal{L}(X, Y)$ is a Banach space with the operator norm and $X^* = \mathcal{L}(X, R)$.

Open Mapping theorem Let T be a continuous linear operator from a Banach space X onto a Banach space Y. Then, T is an open map, i.e., maps every open set to of X onto an open set in Y.

Proof: Suppose $A : X \to Y$ is a surjective continuous linear operator. In order to prove that A is an open map, it is sufficient to show that A maps the open unit ball in X to a neighborhood of the origin of Y. Let $U = B_1(0)$ and $V = B_1(0)$ be a open unit ball of 0 in X and Y, respectively. Then $X = \bigcup_{k \in N} kU$. Since A is surjective,

$$Y = A(X) = A\left(\bigcup_{k\in\mathbb{N}} kU\right) = \bigcup_{k\in\mathbb{N}} A(kU).$$

Since Y is a Banach, by the Baire's category theorem there exists a $k \in N$ such that $\left(\overline{A(kU)}\right)^\circ \neq \varnothing$. Let $c \in \left(\overline{A(kU)}\right)^\circ$. Then, there exists $r > 0$ such that $B_r(c) \subseteq \left(\overline{A(kU)}\right)^\circ$. If $v \in V$, then

$$c,\ c + rv \in B_r(c) \subseteq \left(\overline{A(kU)}\right)^\circ \subseteq \overline{A(kU)}.$$

and the difference $rv = (c + rv) - c$ of elements in $\left(\overline{A(kU)}\right)^\circ$ satisfies $rv \in \subseteq \overline{A(2kU)}$. Since A is linear, $V \subseteq \overline{A\left(\frac{2k}{r}U\right)}$. Thus, it follows that for all $y \in Y$ and $\epsilon > 0$,

$$\text{there exists } x \in X \text{ such that } |x|_X < \tfrac{2k}{r}|y|_Y \text{ and } |y - Ax|_Y < \epsilon. \tag{2.1}$$

Fix $y \in \delta V$, by (2.1), there is some $x_1 \in X$ with $|x_1| < 1$ and $|y - Ax_1| < \frac{\delta}{2}$. By (2.1) one can find a sequence $\{x_n\}$ inductively by

$$|x_n| < 2^{-(n-1)} \quad \text{and} \quad |y - A(x_1 + x_2 + \cdots + x_n)| < \delta 2^{-n}. \tag{2.2}$$

Let $s_n = x_1+x_2+\cdots+x_n$. From the first inequality in (2.2), $\{s_n\}$ is a Cauchy sequence, and since X is complete, $\{s_n\}$ converges to some $x \in X$. By (2.2), the sequence As_n tends to y in Y, and thus $Ax = y$ by continuity of A. Also, we have

$$|x| = \lim_{n\to\infty} |s_n| \leq \sum_{n=1}^{\infty} |x_n| < 2.$$

This shows that every $y \in \delta\, V$ belongs to $A(2U)$, or equivalently, that the image $A(U)$ of $U = B_1(0)$ in X contains the open ball $\frac{\delta}{2}V \subset Y$. □

Corollary (Open Mapping theorem) A bounded linear operator T on Banach spaces is bounded invertible if and only if it is one-to-one and onto.

Definition (Closed Linear Operator) (1) The graph $G(T)$ of a linear operator T on the domain $\mathcal{D}(T) \subset X$ into Y is the set $(x, Tx) : x \in \mathcal{D}(T)\}$ in the product space $X \times Y$. Then T is closed if its graph $G(T)$ is a closed linear subspace of $X \times Y$, i.e.,

if $x_n \in \mathcal{D}(T)$ converges strongly to $x \in X$ and Tx_n converges strongly to $y \in Y$, then $x \in \mathcal{D}(T)$ and $y = Tx$. Thus the notion of a closed linear operator is an extension of the notion of a bounded linear operator.
(2) A linear operator T is said be closable if $x_n \in \mathcal{D}(T)$ converges strongly to 0 and Tx_n converges strongly to $y \in Y$, then $y = 0$.

For a closed linear operator T, the domain $\mathcal{D}(T)$ is a Banach space if it is equipped by the graph norm

$$|x|_{\mathcal{D}(T)} = (|x|_X^2 + |Tx|_Y^2)^{\frac{1}{2}}.$$

Example (Closed linear Operator) Let $T = \frac{d}{dt}$ with $X = Y = L^2(0,1)$ is closed and

$$\begin{aligned} dom\,(A) = H^1(0,1) = \{f \in L^2(0,1) &: \text{absolutely continuous functions on } [0,1] \\ &\text{with square integrable derivative}\}. \end{aligned}$$

If $y_n = Tx_n$, then

$$x_n(t) = x_n(0) + \int_0^t y_n(s)\,ds.$$

If $x_n \in dom(T) \to x$ and $y_n \to y$ in $L^2(0,1)$, then letting $n \to \infty$ we have

$$x(t) = x(0) + \int_0^t y(s)\,ds,$$

i,e., $x \in dom\,(T)$ and $Tx = y$. In general if for $\lambda\,I + T$ for some $\lambda \in R$ has a bounded inverse $(\lambda\,I + T)^{-1}$, then $T : dom\,(A) \subset X \to X$ is closed. In fact, $Tx_n = y_n$ is equivalent to

$$x_n = (\lambda\,I + T)^{-1}(y_n + \lambda\,x_n$$

Suppose $x_n \to x$ and $y_n \to y$ in X, letting $n \to \infty$ in this, we have $x \in dom\,(T)$ and $Tx = T(\lambda\,I + T)^{-1}(\lambda\,x + y) = y$.

Definition (Dual Operator) Let T be a linear operator X into Y with dense domain $\mathcal{D}(T)$. The dual operator of T^* of T is a linear operator on Y^* into X^* defined by

$$\langle y^*, Tx\rangle_{Y^* \times Y} = \langle T^* y^*, x\rangle_{X^* \times X}$$

for all $x \in \mathcal{D}(T)$ and $y^* \in \mathcal{D}(T^*)$.

In fact, for $y^* \in Y^*$ $x^* \in X^*$ satisfying

$$\langle y^*, Tx\rangle = \langle x^*, x\rangle \text{ for all } x \in \mathcal{D}(T)$$

is uniquely defined if and only if $\mathcal{D}(T)$ is dense. The only if part follows since if $\overline{\mathcal{D}(T)} \neq X$ then the Hahn-Banach theory there exits a nonzero $x_0^* \in X^*$ such that $\langle x_0^*, x\rangle = 0$ for all $\mathcal{D}(T)$, which contradicts to the uniqueness assumption. If T is bounded with $\mathcal{D}(T) = X$ then T^* is bounded with $\|T\| = \|T^*\|$.

Examples Consider the gradient operator $T : L^2(\Omega) \to L^2(\Omega)^n$ as

$$Tu = \nabla u = (D_{x_1}u, \cdots D_{x_n}u)$$

with $\mathcal{D}(T) = H^1(\Omega)$. The, we have for $v \in L^2(\Omega)^n$

$$T^* v = -div\,v = -\sum D_{x_k} v_k$$

with domain $\mathcal{D}(T^*) = \{v \in L^2(\Omega)^n : div\, v \in L^2(\Omega) \text{ and } n\cdot v = 0 \text{ at the boundary } \partial\Omega\}$. In fact by the divergence theorem

$$(Tu, v) = \int_\Omega \nabla u \cdot v \int_{\partial\Omega} (n \cdot v) u\, ds - \int_\Omega u(div\, v)\, dx = (u, T^* v)$$

for all $v \in C^1(\Omega)$. First, let $u \in H_0^1(\Omega)$ we have $T^*v = -div\, v \in L^2(\Omega)$ since $H_0^1(\Omega)$ is dense in $L^2(\Omega)$. Thus, $n \cdot v \in L^(\partial\Omega)$ and $n \cdot v = 0$.

Definition (Hilbert space Adjoint operator) Let X, Y be Hilbert spaces and T be a linear operator X into Y with dense domain $\mathcal{D}(T)$. The Hilbert self adjoint operator of T^* of T is a linear operator on Y into X defined by

$$(y, Tx)_Y = (T^*y, x)_X$$

for all $x \in \mathcal{D}(T)$ and $y \in \mathcal{D}(T^*)$. Note that if we let $T' : Y^* \to X^*$ is the dual operator of T, then

$$T^* R_{Y^*\to Y} = R_{X^*\to X} T'$$

where $R_{X^*\to X}$ and $R_{Y^*\to Y}$ are the Riesz maps.

Examples (self-adjoint operator) Let $X = L^2(\Omega)$ and T be the Laplace operator

$$Tu = \Delta u = \sum_{k=1}^{n} D_{x_k x_k} u$$

with domain $\mathcal{D}(T) = H^2(\Omega) \cap H_0^1(\Omega)$. Then T is sel-adjoint, i.e., $T^* = T$. In fact

$$(Tu, v)_X = \int_\Omega \Delta u\, v\, dx = \int_{\partial\Omega} ((n \cdot \nabla u)v - (n \cdot \nabla v)u)\, ds + \int_\Omega \Delta v\, u\, dx = (x, T^*v)$$

for all $v \in C^1(\Omega)$.

Closed Range Theorem Let X and Y be Banach spaces and T a closed linear operator on $\mathcal{D}(T) \subset X$ into Y. Assume that $\mathcal{D}(T)$ is dense in X. The the following properties are all equivalent.
(a) $R(T)$ is closed in Y.
(b) $R(T^*)$ is closed in X^*.
(c) $R(T) = N(T^*)^\perp = \{y \in Y : \langle y^*, y\rangle = 0 \text{ for all } y^* \in N(T^*)\}$.
(d) $R(T^*) = N(T)^\perp = \{x^* \in X^* : \langle x^*, x\rangle = 0 \text{ for all } x^* \in N(T)\}$.

Proof: If $y^* \in N(T^*)^\perp$, then

$$\langle Tx, y^*\rangle = \langle x, T^*y^*\rangle$$

for all $x \in dom(T)$ and $N(T^*)^\perp \subset \overline{R(T)}$. If $y \in R(T)$, then

$$\langle y, y^*\rangle = \langle x, T^*y^*\rangle$$

for all $y^* \in N(T^*)$ and $R(T) \subset N(T^*)^\perp$. Thus, $\overline{R(T)} = N(T^*)^\perp$.

Since T is closed, the graph $G = G(T) = \{(x, Tx) inX \times Y : x \in dom(T)\}$ is closed. Thus, G is a Banach space equipped with norm $|\{x, y\}| = |x| + |y|$ of $X \times Y$. Define a continuous linear operator S from G into Y by

$$S\{x, Tx\} = Tx$$

Then, the dual operator S^* of S is continuous linear operator from Y^* intoG^* and we have

$$\langle \{x, Tx\}, S^*y^*\rangle = \langle S\{x, Tx\}, y^*\rangle = \langle Tx, y^*\rangle = \langle \{x, Tx\}, \{0, y^*\}\rangle$$

for $x \in dom(T)$ and $y^* \in Y^*$. Thus, the functional $\tilde{S}$ $\tilde{S}y^* = S * y^* - \{0, y^*\} : Y^* \to X^* \times Y^*$ vanishes for all elements of G. But, since $\langle \{x, Tx\}, \{x^*, y_1^*\}$ for all $x \in dom(T)$, is equivalent to $\langle x, x^*\rangle = \langle -Tx, y_1^*\rangle$ for all $x \in dom(T)$, i.e., $-T^*y_1^* = x^*$,

$$S^*y^* = \{0, y^*\} + \{-T^*y_1^*, y_1^*\} = \{-T^*y_1^*, y^* + y_1^*\}$$

for all $y^* \in Y^*$. Since y_1^* is arbitrary, $R(S^*) = R(T^*) \times Y^*$. Therefore, $R(S^*)$ is closed in $X^* \times Y^*$ if and only if $R(T)$ is closed and since $R(S) = R(T)$, $R(S)$ is closed if and only if $R(T)$ is closed in Y. Hence we have only prove the equivalency of (a) and (b) in the special case of a bounded operator S.

If T is bounded, then (a) implies(b). Since $R(T)$ a Banach space, by the Hahn-Banach theorem, one can assume $R(T) = Y$ without loss of generality

$$\langle Tx, y_1^*\rangle = \langle x, T^*y_1^*\rangle$$

By the open mapping theorem, there exists $c > 0$ such that for each $y \in Y$ there exists an $x \in X$ with $Tx = y$ and $|x| \leq c\,|y|$. Thus, for $y^* \in Y^*$, we have

$$|\langle y, y^*\rangle| = |\langle Tx, y^*\rangle| = |\langle x, T^*y^*\rangle| \leq c|y||T^*y^*|.$$

and

$$|y^*| \leq \sup_{|y|\leq 1} |\langle y, y^*\rangle| \leq c|T^*y^*|$$

and so $(T^*)^{-1}$ exists and bounded and $D(T^*) = R(T^*)$ is closed.

Let $T_1 : X \to Y_1 = \overline{R(T)}$ be defined by $T_1x = Tx$ For $y_1^* \in Y_1^*$

$$\langle T_1x, y_1^*\rangle = \langle x, T^*y^*\rangle = 0, \quad \text{for all } x \in X$$

and since $R(T)$ is dense in $\overline{R(T)}$, $y_1^* = 0$. Since $R(T^*) = R(T_1^*)$ is closed, $(T_1^*)^{-1}$ exists on $Y_1 = (Y_1)^*$

2.5 Distribution and Generalized Derivatives

In this section we introduce the distribution (generalized function). The concept of distribution is very essential for defining a generalized solution to PDEs and provides the foundation of PDE theory. Let $\mathcal{D}(\Omega)$ be a vector space of all infinitely many continuously differentiable functions $C_0^\infty(\Omega)$ with compact support in Ω. For any compact set K of Ω, let $\mathcal{D}_K(\Omega)$ be the set of all functions $f \in C_0^\infty(\Omega)$ whose support are in K. Define a family of seminorms on $\mathcal{D}(\Omega)$ by

$$p_{K,m}(f) = \sup_{x\in K} \sup_{|s|\leq m} |D^s f(x)|$$

where

$$D^s = \left(\frac{\partial}{\partial x_1}\right)^{s_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{s_n}$$

where $s = (s_1, \cdots, s_n)$ is nonnegative integer valued vector and $|s| = \sum s_k \leq m$. Then, $\mathcal{D}_K(\Omega)$ is a locally convex topological space.

Definition (Distribution) A linear functional T defined on $C_0^\infty(\Omega)$ is a distribution if for every compact subset K of Ω, there exists a positive constant C and a positive integer k such that

$$|T(\phi)| \leq C\, sup_{|s|\leq k,\, x\in K}\, |D^s\phi(x)| \text{ for all } \phi \in \mathcal{D}_K(\Omega).$$

Definition (Generalized Derivative) A distribution S defined by

$$S(\phi) = -T(D_{x_k}\phi) \text{ for all } \phi \in C_0^\infty(\Omega)$$

is called the distributional derivative of T with respect to x_k and we denote $S = D_{x_k}T$.

In general we have

$$S(\phi) = D^sT(\phi) = (-1)^{|s|}\, T(D^s\phi) \text{ for all } \phi \in C_0^\infty(\Omega).$$

This definition is naturally followed from that for f is continuously differentiable

$$\int_\Omega D_{x_k} f\phi\, dx = -\int_\Omega f \frac{\partial}{\partial x_k}\phi\, dx$$

and thus $D_{x_k} f = D_{x_k}T_f = T_{\frac{\partial}{\partial x_k}} f$. Thus, we let $D^s f$ denote the distributional derivative of T_f.

Example (Distribution) (1) For f is a locally integrable function on Ω, one defines the corresponding distribution by

$$T_f(\phi) = \int_\Omega f\phi\, dx \text{ for all } \phi \in C_0^\infty(\Omega).$$

since

$$|T_f(\phi)| \leq \int_K |f|\, dx \sup_{x\in K}\, |\phi(x)|.$$

(2) $T(\phi) = \phi(0)$ defines the Dirac delta δ_0 at $x = 0$, i.e.,

$$|\delta_0(\phi)| \leq \sup_{x\in K}\, |\phi(x)|.$$

(3) Let H be the Heaviside function defined by

$$H(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

Then,

$$D_{T_H}(\phi) = -\int_{-\infty}^{\infty} H(x)\phi'(x)\, dx = \phi(0)$$

and thus $DT_H = \delta_0$ is the Dirac delta function at $x = 0$.

(4) The distributional solution for $-D^2u = \delta_{x_0}$ satisfies

$$-\int_{-\infty}^{\infty} u\phi''\, dx = \phi(x_0)$$

for all $\phi \in C_0^\infty(R)$. That is, $u = \frac{1}{2}|x - x_0|$ is the fundamental solution, i.e.,

$$-\int_{-\infty}^{\infty} |x - x_0|\phi''\,dx = \int_{\infty}^{x_0} \phi'(x)\,dx - \int_{x_0}^{\infty} \phi'(x)\,dx = 2\phi(x_0).$$

In general for $d \geq 2$ let

$$G(x, x_0) = \begin{cases} \frac{1}{4\pi}\log|x - x_0| & d = 2 \\ \\ c_d\,|x - x_0|^{2-d} & d \geq 3. \end{cases}$$

Then

$$\Delta G(x, x_0) = 0, \quad x \neq x_0.$$

and $u = G(x, x_0)$ is the fundamental solution to to $-\Delta$ in R^d,

$$-\Delta u = \delta_{x_0}.$$

In fact, let $B_\epsilon = \{|x - x_0| \leq \epsilon\}$ and $\Gamma = \{|x - x_0| = \epsilon\}$ be the surface. By the divergence theorem

$$\int_{R^d \setminus B_\epsilon(x_0)} G(x, x_0)\Delta\phi(x)\,dx = \int_\Gamma \frac{\partial}{\partial\nu}\phi(G(x, x_0) - \frac{\partial}{\partial\nu}G(x, x_0)\phi(s))\,ds$$

$$= \int_\Gamma (\epsilon^{2-d}\frac{\partial\phi}{\partial\nu} - (2 - d)\epsilon^{1-d}\phi(s))\,ds \to \frac{1}{c_d}\phi(x_0)$$

That is, $G(x, x_0)$ satisfies

$$-\int_{R^d} G(x, x_0)\Delta\phi\,dx = \phi(x_0).$$

In general let $\mathcal{L}$ be a linear diffrenrtial operator and $\mathcal{L}^*$ denote the formal adjoint operator of $\mathcal{L}$ An locally integrable function u is said to be a distributional solution to $\mathcal{L}u = T$ where $\mathcal{L}$ with a distribution T if

$$\int_\Omega u(\mathcal{L}^*\phi)\,dx = T(\phi)$$

for all $\phi \in C_0^\infty(\Omega)$.

Definition (Sovolev space) For $1 \leq p < \infty$ and $m \geq 0$ the Sobolev space is

$$W^{m,p}(\Omega) = \{f \in L^p(\Omega) : D^s f \in L^p(\Omega),\ |s| \leq m\}$$

with norm

$$|f|_{W^{m,p}(\Omega)} = \left(\int_\Omega \sum_{|s| \leq m} |D^s f|^p\,dx\right)^{\frac{1}{p}}.$$

That is,

$$|D^s f(\phi)| \leq c\,|\phi|_{L^q} \text{ with } \frac{1}{p} + \frac{1}{q} = 1.$$

Remark (1) $X = W^{m,p}(\Omega)$ is complete. In fact If $\{f_n\}$ is Cauchy in X, then $\{D^s f_n\}$ is Cauchy in $L^p(\Omega)$ for all $|s| \leq m$. Since $L^p(\Omega)$ is complete, $D^s f_n \to g^s$ in $L^p(\Omega)$. But since

$$\lim_{n\to\infty} \int_\Omega f_n D^s\phi\,dx = \int_\Omega f D^s\phi\,dx = \int g^s \phi\,dx,$$

we have $D^s f = g^s$ for all $|s| \leq m$ and $|f_n - f|_X \to 0$ as $n \to \infty$.
(2) $H^{m,p} \subset W^{1,p}(\Omega)$. Let $H^{m,p}(\Omega)$ be the completion of $C^m(\Omega)$ with respect to $W^{I,p}(\Omega)$ norm. That is, $f \in H^{m,p}(\Omega)$ there exists a sequence $f_n \in C^m(\Omega)$ such that $f_n \to f$ and $D^s f_n \to g^s$ strongly in $L^p(\Omega)$ and thus

$$D^s f_n(\phi) = (-1)^{|s|} \int_\Omega D_s f_n \phi\,dx \to (-1)^{|s|} \int_\Omega g^s \phi\,dx$$

which implies $g^s = D^s f$ and $f \in W^{1,p}(\Omega)$.
(3) If Ω has a Lipschitz continuous boundary, then

$$W^{m,p}(\Omega) = H^{m,p}(\Omega).$$

2.6 Compact Operator

A compact operator is a linear operator A from a Banach space X to another Banach space Y, such that the image under A of any bounded subset of X is a relatively compact subset of Y, i.e., its closure is compact. For example, the integral operator is a concrete example of compact operators. A Fredholm integral equation gives rise to a compact operator A on function spaces.

A compact set in a Banach space is bounded. The converse doses not hold in general. The the closed unit sphere is weakly star compact but not strongly compact unless X is of finite dimension.

Let $C(S)$ be the space of continuous functions on a compact metric space S with norm $|x| = \max_{s\in S} |x(s)|$. the Arzela-Ascoli theorem gives the necessary and sufficient condition for a subset $\{x_\alpha\}$ being (strongly) relatively compact.
Arzela-Ascoli theorem A family $\{x_\alpha\}$ in $C(S)$ is strongly compact if and only if the following conditions hold:

$$\sup_\alpha |x_\alpha| < \infty \text{ (equi-bouned)}$$

$$\lim_{\delta\to 0^+} \sup_{\alpha,\,|s'-s''|\leq\delta} |x_\alpha(s') - x_\alpha(s')| = 0 \text{ (equi-continuous).}$$

Proof: By the Bolzano-Weirstrass theorem for a fixed $s \in S$ $\{x_\alpha(s)\}$ contains a convergent subsequence. Since S is compact, there is a countable dense subset $\{s_n\}$ in S such that for every $\epsilon > 0$ there exists n_ϵ satisfying

$$\sup_{s\in S} \inf_{1\leq j\leq n_\epsilon} dist(s, s_j) \leq \epsilon.$$

Let us denote a convergent subsequence of $\{x_n(s_1)\}$ by $\{x_{n_1}(s_1)\}$. Similarly, the sequence $\{\langle x_{n_1}(s_2)\}$ contains a a convergent subsequence $\{x_{n_2}(s_2)\}$. Continuing in this manner we extract subsequences $\{x_{n_k}(s_k)\}$. The diagonal sequence $\{x^*_{nn}(s)\}$ converges for $s = s_1,\ s_2,\ \cdots$, simultaneously for the dense subset $\{s_k\}$. We will show that $\{x_{nn}\}$

is a Cauchy sequence in $C(S)$. By the equi-continuity of $\{x_n(s)\}$ for all $\epsilon > 0$ there exists δ_ϵ such that $|x_n(s') - x_n(s'')| \leq \epsilon$ for all s', s' satisisfying $dist(s', s'') \leq\leq \delta$. Thus, for $s \in S$ there exists a $j \leq k_\epsilon$ such that

$$|x_{nn}(s) - x_{mm}(s)| \leq |x_{nn}(s) - x_{mm}(s_j)| + |x_{nn}(s_j) - x_{mm}(s_j)| + |x_{mm}(s) - x_{mm}(s_j)|$$

$$\leq 2\epsilon + |x_{nn}(s_j) - x_{mm}(s_j)|$$

which implies that $\lim_{m,m\to\infty} \max_{s\in S} |x_{nn}(s) - x_{mm}(s)| \leq 2\epsilon$ for arbitrary $\epsilon > 0$. $\square$

For the space $L^p(\Omega)$, we have

Frechet-Kolmogorov theorem Let $Omega$ be a subset of R^d. A family $\{x_\alpha\}$ of $L^p(\Omega)$ functions is strongly compact if and only if the following conditions hold:

$$\sup_\alpha \int_\Omega |x_\alpha|^p \, ds < \infty \text{ (equi-bouned)}$$

$$\lim_{|t|\to 0} \sup_\alpha \int_\Omega |x_\alpha(t+s) - x_\alpha(s)|^p \, ds = 0 \text{ (equi-continuous).}$$

Example (Integral Operator) Let $k(x, y)$ be the kernel function satisfying

$$\int_\Omega \int_\Omega |k(x, y)|^2 \, dxdy < \infty$$

Define the integral operator A by

$$Au = \int_\Omega k(x, y)u(y) \, dy \text{ for } u \in X = L^2(\Omega).$$

Then $A \in \mathcal{L}(X)$ i compact. This class of operators A is called the Hilbert-Schmitz operator.

Rellich-Kondrachov theorem Every uniformly bounded sequence in $W^{1,p}(\Omega)$ has a subsequence that converges in $L^q(\Omega)$ provided that $q < p^* = \frac{dp}{d-p}$.

Fredholm Alternative theorem Let $A \in \mathcal{L}(X)$ be compact. For any nonzero $\lambda \in C$, either
1) The equation $Ax - \lambda\, x = 0$ has a nonzero solution x, or
2) The equation $Ax - \lambda, x = f$ has a unique solution x for any function $f \in X$.
In the second case, $(\lambda\, I - A)^{-1}$ is bounded.

Exercise
Problem 1 $T : \mathcal{D}(T) \subset X \to Y$ is closable if and only if the closure of its graph $G(T)$ in $X \times Y$ the graph of a linear operator S.
Problem 2 For a closed linear operator T, the domain $\mathcal{D}(T)$ is a Banach space if it is equipped by the graph norm

$$|x|_{\mathcal{D}(T)} = (|x|_X^2 + |Tx|_Y^2)^{\frac{1}{2}}.$$

3 Constrained Optimization

In this section we develop the Lagrange multiplier theory for the constrained minimization in Banach spaces.

3.1 Hilbert space theory

In this section we develop the Hilbert space theory for the constrained minimization.

Theorem 1 (Riesz Representation) Let X be a Hilbert space. Given $F \in X^*$, consider the minimization

$$J(x) = \frac{1}{2}(x,x)_X - F(x) \quad \text{over } x \in X$$

There exists a unique solution $x^* \in X$ and x^* satisfies

$$(v, x^*) = F(v) \quad \text{for all } v \in X \tag{3.1}$$

The solution $x^* \in X$ is the Riesz representation of $F \in X^*$ and it satisfies $|x^*|_X = |F|_{X^*}$.

Proof: Suppose x^* is a minimizer. Then

$$\frac{1}{2}(x^* + tv, x^* + tv) - F(x^* + tv) \geq \frac{1}{2}(x^*, x^*) - F(x^*)$$

for all $t \in R$ and $v \in X$. Thus, for $t > 0$

$$(v, x^*) - F(v) + \frac{t}{2}|v|^2 \geq 0$$

Letting $t \to 0^+$, we have $(v, x^*) - F(v) \geq 0$ for all $v \in X$. If $v \in X$, then $-v \in X$ and this implies (3.1).

(Uniqueness) Suppose x_1^*, x_2^* are minimizers. Then,

$$(v, .x_1^*) - (v, x_2^*) = F(v) - F(v) = 0 \quad \text{for all } v \in X$$

Letting $v = x_1^* - x_2^*$, we have $|x_1^* - x_2^*|^2 = (x_1^* - x_2^*, x_1^* - x_2^*) = 0$, i.e., $x_1^* = x_2^*$. $\square$

(Existence) Suppose $\{x_n\}$ is a minimizing sequence, i.e., $J(x_n)$ is decreasing and $\lim_{n\to\infty} J(x_n) = \delta = \inf_{x\in X} J(x)$. Note that

$$|x_n - x_m|^2 + |x_n + x_m|^2 = 2(|x_n|^2 + |x_m|^2)$$

Thus,

$$\frac{1}{4}|x_n - x_m|^2 = J(x_m) + J(x_n) - 2J(\frac{x_n + x_m}{2}) - F(x_n) - F(x_m) + 2F(\frac{x_n + x_m}{2}) \leq J(x_n) + J(x_m) - 2\delta.$$

This implies $\{x_n\}$ is a Cauchy sequence in X. Since X is complete there exists $x^* = \lim_{n\to\infty} x_n$ in X and $J(x^*) = \delta$, i.e, x^* is the minimizer. $\square$

Example (Mechanical System)

$$\min \quad \int_0^1 \frac{1}{2}|\frac{d}{dx}u(x)|^2 - \int_0^1 f(x)u(x)\,dx \tag{3.2}$$

over $u \in H_0^1(0,1)$. Here

$$X = H_0^1(0,1) = \{u \in H^1(0,1) : \; u(0) = u(1) = 0\}$$

is a Hilbert space with the inner product

$$(u,v)_{H_0^1} = \int_0^1 \frac{d}{dx}u\frac{d}{dx}v\,dx.$$

$H^1(0,1)$ is a space of absolute continuous functions on $[0,1]$ with square integrable derivative, i.e.,

$$u(x) = u(0) + \int_0^x g(x)\,dx, \quad g(x) = \frac{d}{dx}u(x), \quad \text{a.e.} \quad \text{and } g \in L^2(0,1).$$

The solution $y_f = x^*$ to (3.8) satisfies

$$\int_0^1 \frac{d}{dx}v\frac{d}{dx}y_f\,dx = \int_0^1 f(x)v(x)\,dx.$$

By the integration by part

$$\int_0^1 f(x)v(x)\,dx = \int_0^1 (\int_x^1 f(s)\,ds)\frac{d}{dx}v(x)\,dx,$$

and thus

$$\int_0^1 (\frac{d}{dx}y_f - \int_x^1 f(s)\,ds)\frac{d}{dx}v(x)\,dx = 0$$

for all $v \in H_0^1(0,1)$. This implies that

$$\frac{d}{dx}y_f - \int_x^1 f(s)\,ds = c = \text{a constant}$$

Integrating this, we have

$$y_f = \int_0^x \int_x^1 f(s)\,dsdx + cx$$

Since $y_f(1) = 0$ we have

$$c = -\int_0^1 \int_x^1 f(s)\,dsdx.$$

Moreover, $\frac{d}{dx}y_f \in H^1(0,1)$ and

$$-\frac{d^2}{dx^2}y_f(x) = f(x) \quad \text{a.e.}$$

with $y_f(0) = y_f(1) = 0$.

Now, we consider the case when $F(v) = v(1/2)$. Then,

$$|F(v)| = |\int_0^{1/2} \frac{d}{dx}v\,dx| \leq (\int_0^{1/2} |\frac{d}{dx}v|\,dx)^{1/2}(\int_0^{1/2} 1\,dx)^{1/2} \leq \frac{1}{\sqrt{2}}|v|_X.$$

That is, $F \in X^*$. Thus, we have

$$\int_0^1 (\frac{d}{dx}y_f - \chi_{(0,1/2)}(x))\frac{d}{dx}v(x)\,dx = 0$$

for all $v \in H_0^1(0,1)$, where χ_S is the characteristic function of a set S. It is equivalent to

$$\frac{d}{dx}y_f - \chi_{(0,1/2)}(x) = c = \text{a constant}.$$

Integrating this, we have

$$y_f(x) = \begin{cases} (1+c)x & x \in [0, \frac{1}{2}] \\ \\ cx + \frac{1}{2} & x \in [\frac{1}{2}, 1]. \end{cases}$$

Since $y_f(1) = 0$ we have $c = -\frac{1}{2}$. Moreover y_f satisfies

$$-\frac{d^2}{dx^2}y_f = 0, \;\; \text{for } x \neq \frac{1}{2} \quad y_f(0) = y_f(1) = 0$$

$$y_f((\tfrac{1}{2})^-) = y_f(\tfrac{1}{2})^+), \quad \tfrac{d}{dx}y_f((\tfrac{1}{2})^-) - \tfrac{d}{dx}y_f((\tfrac{1}{2})^+) = 1.$$

Or, equivalently

$$-\frac{d^2}{dx^2}y_f = \delta_{\frac{1}{2}}(x)$$

where δ_{x_0} is the Dirac delta distribution at $x = x_0$.

Theorem 2 (Orthogonal Decomposition $X = \overline{M} \oplus M^\perp$**)** Let X be a Hilbert space and M is a linear subspace of X. Let $\overline{M}$ be the closure of M and $M^\perp = \{z \in X : (z, y) = 0 \text{ for all } y \in M\}$ be the orthogonal complement of M. Then every element $u \in X$ has the unique representation

$$u = u_1 + u_2, \quad u_1 \in \overline{M} \text{ and } u_2 \in M^\perp.$$

Proof: Consider the minimum distance problem

$$\min \quad |x - u|^2 \quad \text{over } x \in \overline{M}. \tag{3.3}$$

As for the proof of Theorem 1 one can prove that there exists a unique solution x^* for problem (3.3) and $x^* \in \overline{M}$ satisfies

$$(x^* - u, y) = 0, \quad \text{for all } y \in \overline{M}$$

i.e., $u - x^* \in M^\perp$. If we let $u_1 = x^* \in \overline{M}$ and $u_2 = u - x^* \in M^\perp$, then $u = u_1 + u_2$ is the desired decomposition. □

Example (Closed subspace) (1) $M = span\{(y_k,\; k \in= \{x \in X; x = \sum_{k\in K} \alpha_k\, y_k\}$ is a subspace of X. If K is finite, M is closed.

(2) Let $X = L^2(-1, 1)$ and $M = \{\text{odd functions}\} = \{f \in X : f(-x) = f(x), \text{ a.e.}\}$ and then M is closed. Let $M_1 = span\{\sin(k\pi t)\}$ and $M_2 = span\{\cos(k\pi t)\}$. $L^2(0,1) = M_1 \oplus M_2$.
(3) Let $E \in \mathcal{L}(X, Y)$. Then $M = N(E) = \{x \in X : Ex = 0\}$ is closed. In fact, if $x_n \in M$ and $x_n \to x$ in X, then $0 = \lim_{n\to\infty} Ex_n = E \lim_{n\to\infty} x_n = Ex$ and thus $x \in N(E) = M$.

Example (Hodge-Decomposition) Let $X = L^2(\Omega)^3$ with where Ω is a bounded open set in R^3 with Lipschitz boundary Γ. Let $M = \{\nabla u,\ u \in H^1(\Omega)\}$ is the gradient field. Then,

$$M^\perp = \{\psi \in X : \nabla \cdot \psi = 0,\ \ n \cdot \psi = 0 \text{ at } \Gamma\}$$

is the divergence free field. Thus, we have the Hodge decomposition

$$X = M \oplus M^\perp = \{\text{grad}\, u,\ u \in H^1(\Omega)\} \oplus \{\text{the divergence free field}\}.$$

In fact $\psi \in M^\perp$

$$\int \nabla\phi \cdot \psi = \int_\Gamma n \cdot \psi\phi\, ds - \int_\Omega \nabla\psi\phi\, dx = 0$$

and $\nabla u \in M,\ u \in H^1(\Omega)/R$ is uniquely determined by

$$\int_\Omega \nabla u \cdot \nabla\phi\, dx = \int_\Omega f \cdot \nabla\phi\, dx$$

for all $\phi \in H^1(\Omega)$.

Theorem 3 (Decomposition) Let $X,\ Y$ be Hilbert spaces and $E \in \mathcal{L}(X,Y)$. Let $E^* \in \mathcal{L}(Y,X)$ is the adjoint of E, i.e., $(Ex,y)_Y = (x,E^*y)_X$ for all $x \in X$ and $y \in Y$. Then $N(E) = R(E^*)^\perp$ and thus from the orthogonal decomposition theorem $X = N(E) \oplus \overline{R(E^*)}$.

Proof: Suppose $x \in N(E)$, then

$$(x, E^*y)_X = (Ex, y)_Y = 0$$

for all $y \in Y$ and thus $N(E) \subset R(E^*)^\perp$. Conversely, $x^* \in R(E^*)^\perp$, i.e.,

$$(x^*, E^*y) = (Ex, y) = 0 \quad \text{for all } y \in Y$$

Thus, $Ex^* = 0$ and $R(E^*)^\perp \subset N(E)$. $\square$

Remark Suppose $R(E)$ is closed in Y, then $R(E^*)$ is closed in X and

$$Y = R(E) \oplus N(E^*), \quad X = R(E^*) \oplus N(E)$$

Thus, we have $b = b_1 + b_2,\ b_1 \in R(E),\ b_2 \in N(E^*)$ for all $b \in Y$ and

$$|Ex - b|^2 = |Ex - b_1|^2 + |b_2|^2.$$

Example ($R(E)$ is not closed) Consider $X = L^2(0,1)$ and define a linear operator $E \in \mathcal{L}(X)$ by

$$(Ef)(t) = \int_0^t f(s)\, ds, \ \ t \in (0,1)$$

Since $R(E) \subset \{\text{absolutely continuous functions on } [0,1]\}$, $\overline{R(E)} = L^2(0,1)$ and $R(E)$ is not a closed subspace of $L^2(0,1)$.

3.2 Minimum norm problem

We consider

$$\min \quad |x|^2 \quad \text{subject to } (x, y_i)_X = c_i,\ 1 \le i \le m. \tag{3.4}$$

Theorem 4 (Minimum Norm) Suppose $\{y_i\}$ are linearly independent. Then there exits a unique solution x^* to (3.4) and $x^* = \sum_{i=1}^m \beta_i y_i$ where $\beta = \{\beta_i\}_{i=1}^m$ satisfies

$$G\beta = c, \quad G_{i,j} = (y_i, y_j).$$

Proof: Let $M = span\{y_i\}_{i=1}^m$. Since $X = M \oplus M^\perp$, it is necessary that $x^* \in M$, i.e.,

$$x^* = \sum_{i=1}^m \beta_i y_i.$$

From the constraint $(x^*, y_i) = c_i,\ 1 \le i \le m$, the condition for $\beta \in R^n$ follows. In fact, $x \in X$ satisfies $(x, y_i) = c_i,\ 1 \le i \le m$ if and only if $x = x^* + z,\ z \in M^\perp$. But $|x|^2 = |x^*|^2 + |z|^2$. $\square$

Remark $\{y_i\}$ is linearly independent if and only if the symmetric matrix $G \in R^{n,n}$ is positive definite since

$$(x, Gx)_{R^n} = |\sum_{k=0}^n x_k\, y_k|_X^2 \quad \text{for all } x \in R^n.$$

Theorem 5 (Minimum Norm) Suppose $E \in \mathcal{L}(X, Y)$. Consider the minimum norm problem

$$\min \quad |x|^2 \quad \text{subject to } Ex = c$$

If $R(E) = Y$, then the minimum norm solution x^* exists and

$$x^* = E^* y, \quad (EE^*)y = c.$$

Proof: Since $R(E^*)$ is closed by the closed range theorem, it follows from the decomposition theorem that $X = N(E) \oplus R(E^*)$. Thus, we $x^* = E^* y$ if there exists $y \in Y$ such that $(EE^*)y = c$. Since $Y = N(E^*) \oplus R(E)$, $N(E^*) = \{0\}$. Since if $EE^* x = 0$, then $(x, EE^* x) = |E^* x| = 0$ and $E^* x = 0$. Thus, $N(EE^*) = N(E^*) = \{0\}$. Since $R(EE^*) = R(E)$ it follows from the open mapping theorem EE^* is bounded invertible and there exits a unique solution to $(EE^*)y = c$. $\square$

Remark If $R(E)$ is closed and $c \in R(E)$, we let $Y = R(E)$ and Theorem 5 is valid. For example we define $E \in \mathcal{L}(X, R^m)$ by

$$Ex = ((y_1, x)_X, \cdots, (y_m, x)_X) \in R^m.$$

If $\{y_i\}_{i=1}^m$ are dependent, then span$\{y_i\}_{i=1}^m$ is a closed subspace of R^m
Example (DC-motor Control) Let $(\theta(t), \omega(t)))$ satisfy

$$\begin{aligned} &\tfrac{d}{dt}\omega(t) + \omega(t) = u(t), \quad \omega(0) = 0 \\ &\tfrac{d}{dt}\theta(t) = \omega(t), \quad \theta(0) = 0 \end{aligned} \tag{3.5}$$

This models a control problem for the dc-motor and the control function $u(t)$ is the current applied to the motor, θ is the angle and ω is the angular velocity. The objective to find a control u on $[0,1]$ such the desired state $(\bar{\theta}, 0)$ is reached at time $t = 1$ and with minimum norm

$$\int_0^1 |u(t)|^2\,dt.$$

Let $X = L^2(0,1)$. Given $u \in X$ we have the unique solution to (3.5) and

$$\omega(1) = \int_0^1 e^{t-1} u(t)\,dt$$

$$\theta(t) = \int_0^1 (1 - e^{t-1}) u(t)\,dt$$

Thus, we can formulate this control problem as problem (3.4) by letting $X = L^2(0,1)$ with the standard inner product and

$$y_1(t) = 1 - e^{t-1}, \quad y_2(t) = e^{t-1} \quad c_1 = \bar{\theta}, \quad c_2 = 0.$$

Example (Linear Control System) Consider the linear control system

$$\frac{d}{dt} x(t) = Ax(t) + Bu(t), \quad x(0) = x_0 \in R^n \tag{3.6}$$

where $x(t) \in R^n$ and $u(t) \in R^m$ are the state and the control function, respectively and system matrices $A \in R^{n\times n}$ and $B \in R^{n\times m}$ are given. Consider a control problem of finding a control that transfers $x(0) = 0$ to a desired state $\bar{x} \in R^n$ at time $T > 0$, i.e., $x(T) = \bar{x}$ with minimum norm. The problem can be formulated as (3.4). First, we have the variation of constants formula:

$$x(t) = e^{At} x_0 + \int_0^t e^{A(t-s)} Bu(s)\,ds$$

where e^{At} is the matrix exponential and satisfies $\frac{d}{dt} e^{At} = Ae^{At} = e^{At} A$. Thus, condition $x(T) = \bar{x}$ is written as

$$\int_0^T e^{A(T-t)} Bu(t)\,ds = \bar{x}$$

Let $X = L^2(0,T;R^m)$ and it then follows from Theorem 4 that the minimum norm solution is given by

$$u^*(t) = B^t E^{A^t(T-t)} \beta, \;\; G\beta = \bar{x},$$

provided that the control Gramean

$$G = \int_0^T e^{A(T-s)} BB^t e^{A^t(T-t)}\,dt \in R^{n\times n}$$

is positive on R^n. If G is positive, then control system (3.6) is controllable and it is equivalent to

$$B^t e^{A^t} x = 0, \;\; t \in [0,T] \text{ implies } x = 0.$$

since

$$(x, Gx) = \int_0^T |B^t e^{A^t t} x|^2\,dt$$

Moreover, G is positive if and only if the Kalman rank condition holds; i.e.,

$$rank\ [B, AB, \cdots A^{n-1}B] = n.$$

In fact, $(x, Gx) = 0$ implies $B^t e^{A^t t} x = 0, t \geq 0$ and thus $B^t (A^t)^k = 0$ for all $k \geq 0$ and thus the claim follows from the Cayley-Hamilton theorem.

Example (Function Interpolation)
(1) Consider the minimum norm on $X = H_0^1(0,1)$:

$$\min \quad \int_0^1 |\frac{d}{dx} u|^2\, dx \quad \text{over } H_0^1(0,1) \tag{3.7}$$

subject to $u(x_i) = c_i,\ 1 \leq i \leq m$, where $0 < x_1 < \cdots < x_m < 1$ are given nodes. Let $X = H_0^1(0,1)$. From the mechanical system example we have the piecewise linear function $y_i(x) \in X$ such that $(u, y_i)_X = u(x_i)$ for all $u \in X$ for each i. From Theorem 3 the solution to (3.7) is given by

$$u^*(x) = \sum_{i=1}^m \beta_i\, y_i(x).$$

That is, u^* is a piecewise linear function with nodes $\{x_i\},\ 1 \leq i \leq m$ and nodal values c_i at x_i, i.e., $\beta_i = \dfrac{c_i}{y_i(x_i)},\ 1 \leq i \leq n$.
(2) Consider the Spline interpolation problem (1.2):

$$\min \quad \int_0^1 \frac{1}{2} |\frac{d^2}{dx^2} u(x)|^2\, dx \quad \text{over all functions } u(x)$$

$$\text{subject to the interpolation conditions } u(x_i) = b_i, \quad u'(x_i) = c_i,\ 1 \leq i \leq m. \tag{3.8}$$

Let $X = H_0^2(0,1)$ be the completion of the pre-Hilbert space $C_0^2[0,1]$ = the space of twice continuously differentiable function with $u(0) = u'(0) = 0$ and $u(1) = u'(1) = 0$ with respect to $H_0^2(0,1)$ inner product;

$$(f, g)_{H_0^2} = \int_0^1 \frac{d^2}{dx^2} f \frac{d^2}{dx^2} g\, dx,$$

i.e.,

$$H_0^2(0,1) = \{u \in L^2(0,1) : \frac{d}{dx} u,\ \frac{d^2}{dx^2} u \in L^2(0,1) \text{ with } u(0) = u'(0) = u(1) = u'(1) = 0\}.$$

Problem (3.8) is a minimum norm problem on $x = H_0^2(0,1)$. For $F_i \in X^*$ defined by $F_i(v) = v(x_i)$ we have

$$(y_i, v)_X - F_i(v) = \int_0^1 (\frac{d^2}{dx^2} y_i + (x - x_i)\chi_{[0,x_i]}) \frac{d^2}{dx^2} v\, dx, \tag{3.9}$$

and for $\tilde{F}_i \in X^*$ defined by $\tilde{F}_i(v) = v'(x_i)$ we have

$$(\tilde{y}_i, v)_X - \tilde{F}_i(v) = \int_0^1 (\frac{d^2}{dx^2} \tilde{y}_i - \chi_{[0,x_i]}) \frac{d^2}{dx^2} v\, dx. \tag{3.10}$$

Thus,

$$\frac{d^2}{dx^2}y_i + (x - x_i)\chi_{[0,x_i]} = a + bx, \quad \frac{d^2}{dx^2}\tilde{y}_i - \chi_{[0,x_i]} = \tilde{a} + \tilde{b}x$$

where $a,\ \tilde{a}\ , b,\ \tilde{b}$ are constants and $y_i = RF_i$ and $\tilde{y}_i = R\tilde{F}_i$ are piecewise cubic functions on $[0,1]$. One can determine $y_i,\ \tilde{y}_i$ using the conditions $y(0) = y'(0) = y(1) = y'(0) = 0$. From Theorem 3 the solution to (3.8) is given by

$$u^*(x) = \sum_{i=1}^m (\beta_i\, y_i(x) + \tilde{\beta}_i\, \tilde{y}_i(x))$$

That is, u^* is a piecewise cubic function with nodes $\{x_i\}$, $1 \le i \le m$ and nodal value c_i and derivative b_i at x_i.

Example (Differential form, Natural Boundary Conditions) (1) Let X be the Hilbert space defined by

$$X = H^1(0,1) = \{u \in L^2(0,1) : \frac{du}{dx} \in L^2(0,1)\}.$$

Given $a,\ c \in C[0,1]$ with $a > 0$ $c \ge 0$ and $\alpha \ge 0$ consider the minimization

$$\min \quad \int_0^1 (\frac{1}{2}(a(x)|\frac{du}{dx}|^2 + c(x)|u|^2) - uf)\,dx + \frac{\alpha}{2}\,|u(1)|^2 - gu(1)$$

over $u \in X = H^1(0,1)$. If u^* is a minimizer, it follows from Theorem 1 that $u^* \in X$ satisfies

$$\int_0^1 (a(x)\frac{du^*}{dx}\frac{dv}{dx} + c(x)\,uv - fv)\,dx + \alpha\, u(1)v(1) - gv(1) = 0 \tag{3.11}$$

for all $v \in X$. By the integration by parts

$$\int_0^1 (-\frac{d}{dx}(a(x)\frac{du^*}{dx}) + c(x)u^*(x) - f(x))v(x)\,dx - a(0)\frac{du^*}{dx}(0)v(0) + (a(1)\frac{du^*}{dx}(1) + \alpha\, u^*(1) - g)v(1) = 0,$$

assuming $a(x)\frac{du^*}{dx} \in H^1(0,1)$. Since $v \in H^1(0,1)$ is arbitrary,

$$\begin{aligned} &-\frac{d}{dx}(a(x)\frac{du^*}{dx}) + c(x)\,u^* = f(x), \quad x \in (0,1) \\ &a(0)\frac{du^*}{dx}(0) = 0, \quad a(1)\frac{du^*}{dx}(1) + \alpha\, u^*(1) = g. \end{aligned} \tag{3.12}$$

(4.3) is the weak form and (3.14) is the strong form of the optimality condition.
(2) Let X be the Hilbert space defined by

$$X = H^2(0,1) = \{u \in L^2(0,1) : \frac{du}{dx},\ \frac{d^2u}{dx} \in L^2(0,1)\}.$$

Given $a,\ b,\ c \in C[0,1]$ with $a > 0$ $b,\ c \ge 0$ and $\alpha \ge 0$ consider the minimization on $X = H^2(0,1)$

$$\min \quad \int_0^1 (\frac{1}{2}(a(x)|\frac{d^2u}{dx^2}|^2 + b(x)|\frac{du}{dx}|^2 + c(x)|u|^2) - uf)\,dx + \frac{\alpha}{2}\,|u(1)|^2 + \frac{\beta}{2}\,|u'(1)|^2 - g_1u(1) - g_2u'(1)$$

over $u \in X = H^2(0,1)$ satisfying $u(0) = 1$. If u^* is a minimizer, it follows from Theorem 1 that $u^* \in X$ satisfies

$$\int_0^1 (a(x)\frac{d^2u^*}{dx^2}\frac{d^2v}{dx^2}+b(x)\frac{du^*}{dx}\frac{dv}{dx}+c(x)\,u^*v-fv)\,dx+\alpha\,u^*(1)v(1)+\beta\,(u^*)'(1)v'(1)-g_1v(1)-g_2v'(1)=0 \tag{3.13}$$

for all $v \in X$ satisfying $v(0) = 0$. By the integration by parts

$$\int_0^1 (\frac{d^2}{dx^2}(a(x)\frac{d^2}{dx^2}u^*) - \frac{d}{dx}(b(x)\frac{d}{dx}u^*) + c(x)u^*(x) - f(x))v(x)\,dx$$

$$+(a(1)\frac{d^2u^*}{dx^2}(1) + \beta\,\frac{du^*}{dx}(1) - g_2)v'(1) - a(0)\frac{d^2u^*}{dx^2}(0)v'(0)$$

$$+(-\frac{d}{dx}(a(x)\frac{d^2u^*}{dx^2} + b(x)\frac{du^*}{dx} + \alpha\,u^* - g_1)(1)v(1) - (-\frac{d}{dx}(a(x)\frac{d^2u^*}{dx^2}) + b(x)\frac{du^*}{dx})(0)v(0) = 0$$

assuming $u^* \in H^4(0,1)$. Since $v \in H^2(0,1)$ satisfying $v(0) = 0$ is arbitrary,

$$\frac{d^2}{dx^2}(a(x)\frac{d^2}{dx^2}u^*) - \frac{d}{dx}(b(x)\frac{du^*}{dx}) + c(x)\,u^* = f(x), \quad x \in (0,1)$$

$$u^*(0) = 1, \quad a(0)\frac{d^2u^*}{dx^2}(0) = 0$$

$$a(1)\frac{d^2u^*}{dx^2}(1) + \beta\,\frac{du^*}{dx}(1) - g_2 = 0, \quad (-\frac{d}{dx}(a(x)\frac{d^2u^*}{dx^2}) + b(x)\frac{du^*}{dx})(1) + \alpha\,u^*(1) - g_1 = 0. \tag{3.14}$$

(4.3) is the weak form and (3.14) is the strong (differential) form of the optimality condition.

Next consider the multi dimensional case. Let Ω be a subdomain in R^d with Lipschtz boundary and Γ_0 is a closed subset of the boundary $\partial\Omega$ with $\Gamma_1 = \partial\Omega \setminus \Gamma_0$. Let Γ is a smooth hypersurface in Ω.

$$H^1_{\Gamma_0}(\Omega) = \{u \in H^1(\Omega) : f(s) = 0,\ s \in \Gamma_0\}$$

with norm $|\nabla u|_{L^2(\Omega)}$. Let Γ is a smooth hyper-surface in Ω and consider the weak form of equation for $u \in H^1_{\Gamma_0}(\Omega)$:

$$\int (\sigma(x)\nabla u \cdot \nabla\phi + u(x)\,(\vec{b}(x)\cdot\nabla\phi) + c(x)\,u(x)\phi(x))\,dx + \int_{\Gamma_1} \alpha(s)u(s)\phi(s)\,ds$$

$$-\int_\Gamma g_0(s)\phi(s)\,ds - \int_\Omega f(x)\phi(x)\,dx - \int_{\Gamma_1} g(s)u(s)\,ds = 0$$

for all $H^1_{\Gamma_0}(\Omega)$. Then the differential form for $u \in H^1_{\Gamma_0}(\Omega)$ is given by

$$\begin{cases} -\nabla\cdot(\sigma(x)\nabla u + \vec{b}u) + c(x)u(x) = f \quad \text{in } \Omega \\ \\ n\cdot(\sigma(x)\nabla u + \vec{b}u) + \alpha\,u = g \quad \text{at } \Gamma_1 \\ \\ [n\cdot(\sigma(x)\nabla u + \vec{b}u)] = g_0 \quad \text{at } \Gamma. \end{cases}$$

By the divergence theorem if u satisfies the strong form, then u is a weak solution. Conversely, letting $\phi \in C_0(\Omega \setminus \Gamma)$ we obtain the first equation, i.e., $\nabla \cdot (\sigma \nabla u + \vec{b} u) \in L^2(\Omega)$ Thus, $n \cdot (\sigma \nabla u + \vec{b} u) \in L^2(\Gamma)$ and the third equation holds. Also, $n \cdot (\sigma \nabla u + \vec{b} u) \in L^2(\Gamma_1)$ and the third equation holds.

3.3 Lax-Milgram Theory and Applications

Let H be a Hilbert space with scalar product (ϕ, ψ) and X be a Hilbert space and $X \subset H$ with continuous dense injection. Let X^* denote the strong dual space of X. H is identified with its dual so that $X \subset H = H^* \subset X^*$ (i.e., H is the pivoting space). The dual product $\langle \phi, \psi \rangle$ on $X^* \times X$ is the continuous extension of the scalar product of H restricted to $H \times X$. This framework is called the Gelfand triple.

Let σ is a bounded coercive bilinear form on $X \times X$. Note that given $x \in X$, $F(y) = \sigma(x, y)$ defines a bounded linear functional on X. Since given $x \in X$, $y \to \sigma(x, y)$ is a bounded linear functional on X, say $x^* \in X^*$. We define a linear operator A from X into X^* by $x^* = Ax$. Equation $\sigma(x, y) = F(y)$ for all $y \in X$ is equivalently written as an equation

$$Ax = F \in X^*.$$

Here,

$$\langle Ax, y \rangle_{X^* \times X} = \sigma(x, y), \quad x,\ y \in X,$$

and thus A is a bounded linear operator. In fact,

$$|Ax|_{X^*} \leq \sup_{|y| \leq 1} |\sigma(x, y)| \leq M\,|x|.$$

Let R be the Riesz operator $X^* \to X$, i.e.,

$$|Rx^*|_X = |x^*| \text{ and } (Rx^*, x)_X = \langle x^*, x \rangle \text{ for all } x \in X,$$

then $\hat{A} = RA$ represents the linear operator $\hat{A} \in \mathcal{L}(X, X)$. Moreover, we define a linear operator $\tilde{A}$ on H by

$$\tilde{A}x = Ax \in H$$

with

$$dom\,(\tilde{A}) = \{x \in X : |\sigma(x, y)| \leq c_x\,|y|_H \text{ for all } y \in X\}.$$

That is, $\tilde{A}$ is a restriction of A on $dom\,(\tilde{A})$. We will use the symbol A for all three linear operators as above in the lecture note and its use should be understood by the underlining context.

Lax-Milgram Theorem Let X be a Hilbert space. Let σ be a (complex-valued) sesquilinear form on $X \times X$ satisfying

$$\sigma(\alpha\, x_1 + \beta\, x_2, y) = \alpha\, \sigma(x_1, y) + \beta\, \sigma(x_2, y)$$

$$\sigma(x, \alpha\, y_1 + \beta\, y_2) = \bar{\alpha}\, \sigma(x, y_1) + \bar{\beta}\, \sigma(x, y_2),$$

$$|\sigma(x, y)| \leq M\,|x||y| \qquad \text{for all } \ x, y \in X \quad \text{(Bounded)}$$

and

$$Re\,\sigma(x, x) \geq \delta\,|x|^2 \ \text{ for all } x \in X \text{ and } \delta > 0 \quad \text{(Coercive)}.$$

Then for each $f \in X^*$ there exist a unique solution $x \in X$ to

$$\sigma(x, y) = \langle f, y \rangle_{X^* \times X} \quad \text{for all } \; y \in X$$

and

$$|x|_X \le \delta^{-1} \, |f|_{X^*}.$$

Proof: Let us define the linear operator S from X^* into X by

$$Sf = x, \quad f \in X^*$$

where $x \in X$ satisfies

$$\sigma(x, y) = \langle f, y \rangle \quad \text{for all } \; y \in X.$$

The operator S is well defined since if $x_1, \; x_2 \in X$ satisfy the above, then $\sigma(x_1 - x_2, y) = 0$ for all $y \in X$ and thus $\delta \, |x_1 - x_2|_X^2 \le Re \, \sigma(x_1 - x_2, x_1 - x_2) = 0$.

Next we show that dom(S) is closed in X^*. Suppose $f_n \in dom(S)$, i.e., there exists $x_n \in X$ satisfying $\sigma(x_n, y) = \langle f_n, y \rangle$ for all $y \in X$ and $f_n \to f$ in X^* as $n \to \infty$. Then

$$\sigma(x_n - x_m, y) = \langle f_n - f_m, y \rangle \quad \text{for all } \; y \in X$$

Setting $y = x_n - x_m$ in this we obtain

$$\delta \, |x_n - x_m|_X^2 \le Re \, \sigma(x_n - x_m, x_n - x_m) \le |f_n - f_m|_{X^*} |x_n - x_m|_X.$$

Thus $\{x_n\}$ is a Cauchy sequence in X and so $x_n \to x$ for some $x \in X$ as $n \to \infty$. Since σ and the dual product are continuous, thus $x = Sf$.

Now we prove that dom$(S) = X^*$. Suppose dom$(S) \neq X^*$. Since dom(S) is closed there exists a nontrivial $x_0 \in X$ such that $\langle f, x_0 \rangle = 0$ for all $f \in$dom(S). Consider the linear functional $F(y) = \sigma(x_0, y), \; y \in X$. Then since σ is bounded $F \in X^*$ and $x_0 = SF$. Thus $F(x_0) = 0$. But since $\sigma(x_0, x_0) = \langle F, x_0 \rangle = 0$, by the coercivity of σ $x_0 = 0$, which is a contradiction. Hence dom$(S) = X^*$. $\square$

Assume that σ is coercive. By the Lax-Milgram theorem A has a bounded inverse $S = A^{-1}$. Thus,

$$dom\,(\tilde{A}) = A^{-1} H.$$

Moreover $\tilde{A}$ is closed. In fact, if

$$x_n \in dom\,(\tilde{A}) \to x \text{ and } f_n = A x_n \to f \text{ in } H,$$

then since $x_n = S f_n$ and S is bounded, $x = Sf$ and thus $x \in dom\,(\tilde{A})$ and $\tilde{A} x = f$.

If σ is symmetric, $\sigma(x, y) = (x, y)_X$ defines an inner product on X. and SF coincides with the Riesz representation of $F \in X^*$. Moreover,

$$\langle Ax, y \rangle = \langle Ay, x \rangle \text{ for all } x, y \in X.$$

and thus $\tilde{A}$ is a self-adjoint operator in H.

Example (Laplace operator) Consider $X = H_0^1(\Omega)$, $H = L^2(\Omega)$ and

$$\sigma(u, \phi) = (u, \phi)_X = \int_\Omega \nabla u \cdot \nabla \phi \, dx.$$

Then,

$$Au = -\Delta u = -(\frac{\partial^2}{\partial x_1^2}u + \frac{\partial^2}{\partial x_2^2}u)$$

and

$$dom\,(\tilde{A}) = H^2(\Omega) \cap H_0^1(\Omega).$$

for Ω with C^1 boundary or convex domain Ω.

For $\Omega = (0,1)$ and $f \in L^2(0,1)$

$$\int_0^1 \frac{d}{dx}y\frac{d}{dx}u\,dt = \int_0^1 f(x)y(x)\,dx$$

is equivalent to

$$\int_0^1 \frac{d}{dx}y\,(\frac{d}{dx}u + \int_x^1 f(s)\,ds)\,dx = 0$$

for all $y \in H_0^1(0,1)$. Thus,

$$\frac{d}{dx}u + \int_x^1 f(s)\,ds = c \text{ (a constant)}$$

and therefore $\frac{d}{dx}u \in H^1(0,1)$ and

$$Au = -\frac{d^2}{dx^2}u = f \text{ in } L^2(0,1).$$

Example (Elliptic operator) Consider a second order elliptic equation

$$\mathcal{A}u = -\nabla\cdot(a(x)\nabla u) + b(x)\cdot\nabla u + c(x)u(x) = f(x), \quad \frac{\partial u}{\partial\nu} = g \text{ at } \Gamma_1 \quad u = 0 \text{ at } \Gamma_0$$

where Γ_0 and Γ_1 are disjoint and $\Gamma_0 \cup \Gamma_1 = \Gamma$. Integrating this against a test function ϕ, we have

$$\int_\Omega \mathcal{A}u\phi\,dx = \int_\Omega (a(x)\nabla u\cdot\nabla\phi + b(x)\cdot\nabla u\phi + c(x)u\phi)\,dx - \int_{\Gamma_1} g\phi\,ds_x = \int_\Omega f(x)\phi(x)\,dx,$$

for all $\phi \in C^1(\Omega)$ vanishing at Γ_0. Let $X = H^1_{\Gamma_0}(\Omega)$ is the completion of $C^1(\Omega)$ vanishing at Γ_0 with inner product

$$(u,\phi) = \int_\Omega \nabla u\cdot\nabla\phi\,dx$$

i.e.,

$$H^1_{\Gamma_0}(\Omega) = \{u \in H^1(\Omega) : u|_{\Gamma_0} = 0\}$$

Define the bilinear form σ on $X \times X$ by

$$\sigma(u,\phi) = \int_\Omega (a(x)\nabla u\cdot\nabla\phi + b(x)\cdot\nabla u\phi + c(x)u\phi.$$

Then, by the Green's formula

$$\sigma(u,u) = \int_\Omega (a(x)|\nabla u|^2 + b(x)\cdot\nabla(\frac{1}{2}|u|^2) + c(x)|u|^2)\,dx$$

$$= \int_\Omega (a(x)|\nabla u|^2 + (c(x) - \frac{1}{2}\nabla\cdot b)\,|u|^2)\,dx + + \int_{\Gamma_1} \frac{1}{2}n\cdot b|u|^2\,ds_x.$$

If we assume

$$0 < \underline{a} \leq a(x) \leq \bar{a}, \quad c(x) - \frac{1}{2}\nabla \cdot b \geq 0, \quad n \cdot b \geq 0 \text{ at } \Gamma_1,$$

then σ is bounded and coercive with $\delta = \underline{a}$.

The Banach space version of Lax-Milgram theorem is as follows.

Banach-Necas-Babuska Theorem Let V and W be Banach spaces. Consider the linear equation for $u \in W$

$$a(u,v) = f(v) \quad \text{for all } v \in V \tag{3.15}$$

for given $f \in V^*$, where a is a bounded bilinear form on $W \times V$. The problem is well-posed in if and only if the following conditions hold:

$$\begin{aligned} &\inf_{u\in W} \sup_{v\in V} \frac{a(u,v)}{|u|_W |v|_V} \geq \delta > 0 \\ &a(u,v) = 0 \text{ for all } u \in W \text{ implies } v = 0 \end{aligned} \tag{3.16}$$

Under conditions we have the unique solution $u \in W$ to (3.15) satisfies

$$|u|_W \leq \frac{1}{\delta}\,|f|_{V^*}.$$

Proof: Let A be a bounded linear operator from W to V^* defined by

$$\langle Au, v\rangle = a(u,v) \text{ for all } u \in W,\ v \in V.$$

The inf-sup condition is equivalent to for any w?W

$$|Aw|_{V*} \geq \delta |u|_W,$$

and thus the range of A, $R(A)$, is closed in V^* and $N(A) = 0$. But since V is reflexive and

$$\langle Au, v\rangle_{V^*\times V} = \langle u, A^*v\rangle_{W\times W^*}$$

from the second condition $N(A^*) = \{0\}$. It thus follows from the closed range and open mapping theorems that A^{-1} is bounded. $\square$

Next, we consider the generalized Stokes system. Let V and Q be Hilbert spaces. We consider the mixed variational problem for $(u,p) \in V \times Q$ of the form

$$a(u,v) + b(p,v) = f(v), \quad b(u,q) = g(q) \tag{3.17}$$

for all $v \in V$ and $q \in Q$, where a and b is bounded bilinear form on $V \times V$ and $V \times Q$. If we define the linear operators $A \in \mathcal{L}(V,V^*)$ and $B \in \mathcal{L}(V,Q^*)$ by

$$\langle Au, v\rangle = a(u,v) \quad \text{and} \quad \langle Bu, q\rangle = b(u,q)$$

then it is equivalent to the operator form:

$$\begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}.$$

Assume the coercivity on a

$$a(u,u) \geq \delta\,|u|_V^2 \tag{3.18}$$

and the inf-sup condition on b

$$\inf_{q\in P}\sup_{u\in V}\frac{b(u,q)}{|u|_V|q|_Q} \geq \beta > 0 \tag{3.19}$$

Note that inf-sup condition that for all q there exists $u \in V$ such that $Bu = q$ and $|u|_V \leq \frac{1}{\beta}|q|_Q$. Also, it is equivalent to $|B^*p|_{V^*} \geq \beta\,|p|_Q$ for all $p \in Q$.

Theorem (Mixed problem) Under conditions (3.18)-(3.19) there exits a unique solution $(u,p) \in V \times Q$ to (3.17) and

$$|u|_V + |p|_Q \leq c\,(|f|_{V^*} + |g|_{Q^*})$$

Proof: For $\epsilon > 0$ consider the penalized problem

$$\begin{aligned} &a(u_\epsilon, v) + b(v, P_\epsilon) = f(v), \quad \text{for all } v \in V \\ &-b(u_\epsilon, q) + \epsilon(p_\epsilon, q)_Q = -g(q) \quad \text{for all } q \in Q. \end{aligned} \tag{3.20}$$

By the Lax-Milgram theorem for every $\epsilon > 0$ there exists a unique solution. From the first equation

$$\beta\,|p_\epsilon|_Q \leq |f - Au_\epsilon|_{V^*} \leq |f|_{V^*} + M\,|u_\epsilon|_Q.$$

Letting $v = u_\epsilon$ and $q = p_\epsilon$ in the first and second equation, we have

$$\delta\,|u_\epsilon|_V^2 + \epsilon\,|p_\epsilon|_Q^2 \leq |f|_{V^*}||u_\epsilon|_V + |p_\epsilon|_Q|g|_{Q^*} \leq C\,(|f|_{V^*} + |g|_{Q^*})|u_\epsilon|_V),$$

and thus $|u_\epsilon|_V$ and $|p_\epsilon|_Q$ as well, are bounded uniformly in $\epsilon > 0$. Thus, (u_ϵ, p_ϵ) has a weakly convergent subspace to (u,p) in $V \times Q$ and (u,p) satisfies (3.17). $\square$

3.4 Quadratic Constrained Optimization

In this section we discuss the constrained minimization. First, we consider the case of equality constraint. Let A be coercive and self-adjoint operator on X and $E \in \mathcal{L}(X,Y)$. Let $\sigma : X \times X \to R$ is a symmetric, bounded, ceoercive bilinear form, i.e.,

$$\sigma(\alpha\,x_1 + \beta\,x_2, y) = \alpha\,\sigma(x_1, y) + \beta\,\sigma(x_2, y) \text{ for all } x_1,\ x_2,\ y \in X \text{ and } \alpha,\ \beta \in R$$

$$\sigma(x,y) = \sigma(y,x), \quad |\sigma(x,y)| \leq M\,|x|_X|y|_X, \quad \sigma(x,x) \geq \delta\,|x|^2$$

for some $0 < \delta \leq M < \infty$. Thus, $\sigma(x,y)$ defines an inner product on X. In fact, since σ is bounded and coercive, $\sqrt{\sigma(x,x)} = \sqrt{(Ax,x)}$ is an equivalent norm of X. Since $y \to \sigma(x,y)$ is a bounded linear functional on X, say x^* and define a linear operator from X to X^* by $x^* = Ax$. Then, $A \in \mathcal{L}(X, X^*)$ with $|A|_{X\to X^*} \leq M$ and $\langle Ax, y\rangle_{X^*\times X} = \sigma(x,y)$ for all $x,\ y \in X$. Also, if $R_{X^*\to X}$ is the Riesz map of X^* and define $\hat{A} = R_{X^*\to X}A \in \mathcal{L}X, X)$, then

$$(\hat{A}x, y)_X = \sigma(x,y) \text{ for all } x,\ y \in X,$$

and thus $\hat{A}$ is a self-adjoint operator in X. Throughout this section, we assume $X = X^*$ and $Y = Y^*$, i.e., we identify A as the self-adjoint operator $\tilde{A}$. Consider the unconstrained minimization problem: for $f \in X^*$

$$\min \quad F(x) = \frac{1}{2}\sigma(x,x) - f(x). \tag{3.21}$$

Since for $x,\ v \in X$ and $t \in R$

$$\frac{F(x+tv) - F(x)}{t} = \sigma(x,v) - f(v) + \frac{t}{2}\sigma(v,v),$$

$x \in X \to F(x) \in R$ is differentiable and we have the necessary optimality for (3.21):

$$\sigma(x,v) - f(v) = 0 \text{ for all } v \in X. \tag{3.22}$$

Or, equivalently

$$Ax = f \text{ in } X^*$$

Also, if $R_{X^* \to X}$ is the Riesz map of X^* and define $\tilde{A} = R_{X^* \to X} A \in \mathcal{L}(X,X)$, then

$$(\tilde{A}x, y)_X = \sigma(x,y) \text{ for all } x,\ y \in X,$$

and thus $\tilde{A}$ is a self-adjoint operator in X. Throughout this section, we assume $X = X^*$ and $Y = Y^*$, i.e., we identify A as the self-adjoint operator $\tilde{A}$.

Consider the equality constraint minimization;

$$\min \quad \frac{1}{2}(Ax,x)_X - (a,x)_X \quad \text{subject to } Ex = b \in Y, \tag{3.23}$$

or equivalently for $f \in X^*$

$$\min \quad \frac{1}{2}\sigma(x,x) - f(x) \quad \text{subject to } Ex = b \in Y.$$

The necessary optimality for (3.23) is:

Theorem 6 (Equality Constraint) Suppose there exists a $\bar{x} \in X$ such that $E\bar{x} = b$ then (3.23) has a unique solution $x^* \in X$ and $Ax^* - a \perp N(E)$. If $R(E^*)$ is closed, then there exits a Lagrange multiplier $\lambda \in Y$ such that

$$Ax^* + E^*\lambda = a, \quad Ex^* = b. \tag{3.24}$$

If $range(E) = Y$ (E is surjective), there exits a Lagrange multiplier $\lambda \in Y$ is unique.

Proof: Suppose there exists a $\bar{x} \in X$ such that $E\bar{x} = b$ then (3.23) is equivalent to

$$\min \quad \frac{1}{2}\left(A(z+\bar{x}), z+\bar{x}\right) - (a, z+\bar{x}) \quad \text{over } z \in N(E).$$

Thus, it follows from Theorem 2 that it has a unique minimizer $z^* \in N(E)$ and $(A(z^*+\bar{x}) - a, v) = 0$ for all $v \in N(E)$ and thus $A(z^*+\bar{x}) - a \perp N(E)$. From the orthogonal decomposition theorem $Ax^* - a \in \overline{R(E^*)}$. Moreover if $R(E^*)$ is surjective, from the closed range theorem $\overline{R(E^*)} = R(E^*)$ and there exist exists a $\lambda \in Y$ satisfying

$Ax^* + E^*\lambda = a$. If E is surjective ($R(E) = Y$), then $N(E^*) = 0$ and thus the multiplier λ is unique. $\square$

Remark (Lagrange Formulation) Define the Lagrange functional

$$L(x, \lambda) = J(x) + (\lambda, Ex - b).$$

The optimality condition (3.24) is equivalent to

$$\frac{\partial}{\partial x} L(x^*, \lambda) = 0, \quad \frac{\partial}{\partial \lambda} L(x^*, \lambda) = 0.$$

Here, $x \in X$ is the primal variable and the Lagrange multiplier $\lambda \in Y$ is the dual variable.

Remark (1) If E is not surjective, one can seek a solution using the penalty formulation:

$$\min \quad \frac{1}{2}(Ax, x)_X - (a, x)_X + \frac{\beta}{2}|Ex - b|_Y^2 \tag{3.25}$$

for $\beta > 0$. The necessary optimality condition is

$$Ax_\beta + \beta\, E^*(Ex_\beta - b) - a = 0$$

If we let $\lambda_\beta = \beta\,(Ex - b)$, then we have

$$\begin{pmatrix} A & E^* \\ E & -\frac{I}{\beta} \end{pmatrix} \begin{pmatrix} x_\beta \\ \lambda_\beta \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

the optimality system (3.24) is equivalent to the case when $\beta = \infty$. One can prove the monotonicity

$$J(x_\beta) \uparrow, \quad |Ex_\beta - b|_Y \downarrow \text{ as } \beta \to \infty$$

In fact for $\beta > \hat{\beta}$ we have

$$J(x_\beta) + \tfrac{\beta}{2}|Ex_\beta - b|_Y^2 \leq J(x_{\hat{\beta}}) + \tfrac{\beta}{2}|Ex_{\hat{\beta}} - b|_Y^2$$

$$J(x_{\hat{\beta}}) + \tfrac{\hat{\beta}}{2}|Ex_{\hat{\beta}} - b|_Y^2 \leq J(x_\beta) + \tfrac{\hat{\beta}}{2}|Ex_\beta - b|_Y^2.$$

Adding these inequalities, we obtain

$$(\beta - \hat{\beta})(|Ex_\beta - b|_Y^2 - |Ex_{\hat{\beta}} - b|_Y^2) \leq 0,$$

which implies the claim. Suppose there exists a $\bar{x} \in X$ such that $E\bar{x} = b$ then from Theorem 6

$$J(x_\beta) + \frac{\beta}{2}\,|Ex_\beta - b|^2 \leq J(x^*).$$

and thus $|x_\beta|_X$ is bounded uniformly in $\beta > 0$ and

$$\lim_{\beta \to \infty} |Ex_\beta - b| = 0$$

Since X is a Hilbert space, there exists a weak limit $x \in X$ with $J(x) \leq J(x^*)$ and $Ex = b$ since norms are weakly sequentially lower semi-continuous. Since the optimal solution x^* to (3.23) is unique, $\lim x_\beta = x^*$ as $\beta \to \infty$. Suppose λ_β is bounded in Y, then there exists a subsequence of λ_β that converges weakly to λ in Y. Then, (x^*,λ) satisfies $Ax^*+E^*\lambda=a$. Conversely, if there is a Lagrange multiplier $\lambda^*\in Y$, then λ_β is bounded in Y. In fact, we have

$$(A(x_\beta-x^*),x_\beta-x^*)+(x_\beta-x^*,E^*(\lambda_\beta-\lambda^*))=0$$

where

$$\beta\,(x_\beta-x^*,E^*(\lambda_\beta-\lambda^*))=\beta\,(Ex_\beta-b,\lambda_\beta-\lambda^*)=|\lambda_\beta|^2-(\lambda_\beta,\lambda^*).$$

Thus we obtain

$$\beta\,(x_\beta-x^*,A(x_\beta-x^*))+|\lambda_\beta|^2=(\lambda_\beta,\lambda^*)$$

and hence $|\lambda_\beta|\le|\lambda^*|$. Moreover, if $R(E)=Y$, then $N(E^*)=\{0\}$. Since λ_β satisfies $E^*\lambda_\beta=a-Ax_\beta\in X$, we have

$$\lambda_\beta=Es_\beta,\quad s_\beta=(E^*E)^{-1}(a-Ax_\beta)$$

and thus $\lambda_\beta\in Y$ is bounded uniformly in $\beta>0$.
(2) If E is not surjective but $R(E)$ is closed. We have $\lambda\in R(E)\subset Y$ such that

$$Ax^*+E^*\lambda=a$$

by letting $Y=R(E)$ is a Hilbert space. Also, we have $(x_\beta,\lambda_\beta)\to(x,\lambda)$ satisfying

$$Ax+E^*\lambda=a$$

$$Ex=P_{R(E)}b$$

since

$$|Ex-b|^2=|Ex-P_{R(E)}b|^2+|P_{N(E^*)}b|^2.$$

<u>Example (Parameterized constrained problem)</u>

$$\min\quad \tfrac{1}{2}\,((Ay,y)+\alpha\,(p,p))-(a,y)$$

$$\text{subject to}\quad E_1y+E_2p=b$$

where $x=(y,p)\in X_1\times X_2=X$ and $E_1\in\mathcal{L}(X_1,Y)$, $E_2\in\mathcal{L}(X_2,Y)$. Assume $Y=X_1$ and E_1 is bounded invertible. Then we have the optimality

$$\begin{cases} Ay^*+E_1^*\lambda=a\\[2mm] \alpha\,p+E_2^*\lambda=0\\[2mm] E_1y^*+E_2p^*=b\end{cases}$$

<u>Example (Stokes system)</u> Let $X=H_0^1(\Omega)\times H_0^1(\Omega)$ where Ω is a bounded open set in R^2. A Hilbert space $H_0^1(\Omega)$ is the completion of $C_0(\Omega)$ with respect to the $H_0^1(\Omega)$ inner product

$$(f,g)_{H_0^1(\Omega)}=\int_\Omega \nabla f\cdot\nabla g\,dx$$

Consider the constrained minimization

$$\int_\Omega (\frac{1}{2}(|\nabla u_1|^2 + |\nabla u_2|^2) - f\cdot u)\,dx$$

over the velocity field $u = (u_1, u_2) \in X$ satisfying

$$\operatorname{div} u = \frac{\partial}{\partial x_1}u_1 + \frac{\partial}{\partial x_2}u_2 = 0.$$

From Theorem 5 with $Y = L^2(\Omega)$ and $Eu = \operatorname{div} u$ on X we have the necessary optimality

$$-\Delta u + \operatorname{grad} p = f, \quad \operatorname{div} u = 0$$

which is called the Stokes system. Here $f \in L^2(\Omega)\times L^2(\Omega)$ is a applied force $p = -\lambda^* \in L^2(\Omega)$ is the pressure.

3.5 Variational Inequalities

Let Z be a Hilbert lattice with order $\leq$ and $Z^* = Z$ and $G \in \mathcal{L}(X, Z)$. For example, $Z = L^2(\Omega)$ with a.e. pointwise inequality constarint. Consider the quadratic programming:

$$\min \frac{1}{2}(Ax, x)_X - (a, x)_X \quad \text{subject to } Ex = b, \quad Gx \leq c. \tag{3.26}$$

Note that the constrain set $\mathcal{C} = \{x \in X, Ex = b,\ Gx \leq c\}$ is a closed convex set in X. In general we consider the constrained problem on a Hilbert space X:

$$F_0(x) + F_1(x) \quad \text{over } \ x \in \mathcal{C} \tag{3.27}$$

where $\mathcal{C}$ is a closed convex set, $F : X \to R$ is C^1 and $F_1 : X \to R$ is convex, i.e.

$$F_1((1-\lambda)x_1 + \lambda\, x_2) \geq (1-\lambda)\, F_1(x_1) + \lambda\, F(x_2) \text{ for all } x_, \ x_2 \in X \text{ and } 0 \leq \lambda \leq 1.$$

Then, we have the necessary optimality in the form of the variational inequality:

Theorem 7 (Variational Inequality) Suppose $x^* \in \mathcal{C}$ minimizes (3.27), then we have the optimality condition

$$F_0'(x^*)(x - x^*) + F_1(x) - F_1(x^*) \geq 0 \text{ for all } x \in \mathcal{C}. \tag{3.28}$$

Proof: Let $x_t = x^* + t\,(x - x^*)$ for all $x \in \mathcal{C}$ and $0 \leq t \leq 1$. Since $\mathcal{C}$ is convex, $x_t \in \mathcal{C}$. Since $x^* \in \mathcal{C}$ is optimal,

$$F_0(x_t) - F_0(x^*) + F_1(x_t) - F_1(x^*) \geq 0, \tag{3.29}$$

where

$$F_1(x_t) \leq (1-t)F_1(x^*) + tF(x)$$

and thus

$$F_1(x_t) - F_1(x^*) \leq t\,(F_1(x) - F_1(x^*)).$$

Since

$$\frac{F_0(x_t) - F_0(x^*)}{t} \to F_0'(x^*)(x - x^*) \quad \text{as } \ t \to 0^+,$$

the variational inequality (3.28) follows from (3.29) by letting $t \to 0^+$. $\square$

Example (Inequality constraint) For $X = R^n$

$$F'(x^*)(x - x^*) \geq 0, \quad x \in \mathcal{C}$$

implies that

$$F'(x^*) \cdot \nu \leq 0 \text{ and } F'(x^*) \cdot \tau = 0,$$

where ν is the outward normal and τ is the tangents of $\mathcal{C}$ at x^*. Let $X = R^2$, $\mathcal{C} = [0,1] \times [0,1]$ and assume $x^* = [1, s],\ 0 < s < 1$. Then,

$$\frac{\partial F}{\partial x_1}(x^*) \leq 0, \quad \frac{\partial F}{\partial x_2}(x^*) = 0.$$

If $x^* = [1, 1]$ then

$$\frac{\partial F}{\partial x_1}(x^*) \leq 0, \quad \frac{\partial F}{\partial x_2}(x^*) \leq 0.$$

Example (L^1-optimization) Let U be a closed convex subset in R. Consider the minimization problem on $X = L^2(\Omega)$ and $\mathcal{C} = \{u \in U \text{ a.e. } \ \text{in } \Omega\}$;

$$\min \quad \frac{1}{2}|Eu - b|_Y^2 + \alpha \int_\Omega |u(x)|\, dx. \quad \text{over } u \in \mathcal{C}. \tag{3.30}$$

Then, the optimality condition is

$$(Eu^* - b, E(u - u^*))_Y + \alpha \int_\Omega (|u| - |u^*|)\, dx \geq 0 \tag{3.31}$$

for all $u \in \mathcal{C}$. Let $p = -E^*(Eu^* - b) \in X$ and

$$u = v \in U \text{ on } |x - \bar{x}| \leq \delta, \quad \text{otherwise } u = u^*.$$

From (4.24) we have

$$\frac{1}{\delta} \int_{|x - \bar{x}| \leq \delta} (-p(v - u^*(x)) + \alpha\, (|v| - |u^*(x)|)\, dx \geq 0.$$

Suppose $u^* \in L^1(\Omega)$ and letting $\delta \to 0^+$ we obtain

$$-p(v - u^*(\bar{x})) + \alpha\, (|v| - |u^*(\bar{x})|) \geq 0$$

a.e. $\bar{x} \in \Omega$ (Lebesgue points of $|u^*|$ and for all $v \in U$. That is, $u^* \in U$ minimizes

$$-pu + \alpha\, |u| \quad \text{over } u \in U. \tag{3.32}$$

Suppose $U = R$ it implies that if $|p| < \alpha$, then $u^* = 0$ and otherwise $p = \alpha \frac{u^*}{|u^*|} = \alpha\, sign(u^*)$. Thus, we obtain the necessary optimality

$$(E^*(Eu^* - b))(x) + \alpha\, \partial |u|(u^*(x)) = 0, \quad \text{a.e. in } \ \Omega,$$

where the sub-differential of $|u|$ (see, Section) is defined by

$$\partial|u|(s) = \begin{cases} 1 & s > 0 \\ [-1,1] & s = 0 \\ -1 & s < 0. \end{cases}$$

Theorem 8 (Quadratic Programming) Consider the quadratic programming (3.26):

$$\min \quad \frac{1}{2}(Ax,x)_X - (a,x) \text{ subject to } Ex = b, \;\; Gx \le c.$$

If there exists $\bar{x} \in X$ such that $E\bar{x} = b$ and $G\bar{x} \le c$, then there exits unique solution to (3.26) and

$$(Ax^* - a, x - x^*) \ge 0 \quad \text{for all } x \in \mathcal{C}, \tag{3.33}$$

where $\mathcal{C} = \{Ex = b,\; Gx \le c\}$. Moreover, if $R(E) = Y$ and assume the regular point condition

$$0 \in \text{int}\,\{G\,N(E) - Z^- + G(x^*) - c\},$$

then there exists $\lambda \in Y$ and $\mu \in Z^+$ such that

$$\begin{cases} Ax^* + E^*\lambda + G^*\mu = a \\ \\ Ex^* = b \\ \\ Gx^* - c \le 0, \quad \mu \ge 0 \quad \text{and} \quad (Gx^* - c, \mu)_Z = 0. \end{cases} \tag{3.34}$$

Proof: First note that (3.26) is equivalent to

$$\min \quad \frac{1}{2}(A(z+\bar{x}, z+\bar{x})_X - (a, z+\bar{x})_X \quad \text{subject to } Ez = 0, \quad Gz \le c - G\bar{x} = \hat{c}$$

Since $\widehat{\mathcal{C}} = \{Ez = 0,\; Gz \le \hat{c}\}$ is a closed convex set in the Hilbert space $N(E)$. As in the proof of Theorem 1 the minimizing sequence $z_n \in \widehat{\mathcal{C}}$ satisfies

$$\frac{1}{4}|z_m - z_n| \le J(z_n) + J(z_m) - 2\,\delta \to 0$$

as $m \ge n \to \infty$. Since $\widehat{\mathcal{C}}$ is closed, there exists a unique minimizer $z^* \in \widehat{\mathcal{C}}$ and z^* satisfies

$$(A(z^* + \bar{x}) - a, z - z^*) \ge 0 \quad \text{for all } z \in \widehat{\mathcal{C}},$$

which is equivalent to (3.33). Moreover, it follows from the Lagrange Multiplier theory (below) if the regular point condition holds, then we obtain (3.34). $\square$

Example Suppose $Z = R^m$. If $(Gx^* - c)_i = 0$, then i is called an active index and let G_a the (row) vectors of G corresponding to the active indices. Suppose

$$\begin{pmatrix} E \\ G_a \end{pmatrix} \text{ is surjective (Kuhn-Tuker condition),}$$

then the regular point condition holds.

<u>**Remark (Lagrange Formulation)**</u> Define the Lagrange functional;

$$L(x, \lambda, \mu) = J(x) + (\lambda, Ex - b) + (\mu, (Gx - c)^+).$$

for $(x, \lambda) \in X \times Y,\ \mu \in Z^+$; The optimality condition (3.34) is equivalent to

$$\frac{\partial}{\partial x} L(x^*, \lambda, \mu) = 0, \quad \frac{\partial}{\partial \lambda} L(x^*, \lambda, \mu) = 0, \quad \frac{\partial}{\partial \mu} L(x^*, \lambda, \mu) = 0$$

Here $x \in X$ is the primal variable and the Lagrange multiplier $\lambda \in Y$ and $\mu \in Z^+$ are the dual variables.

<u>**Remark**</u> Let $A \in \mathcal{L}(X, X^*)$ and $E : X \to Y$ is surjective. If we define G is natural injection X to Z. Define $\mu \in X^*$ by

$$-\mu = Ax^* - a + E^*\lambda.$$

Since $\langle t(x^* - c) - (x^* - c), \mu \rangle = (t - 1) \langle x^* - c, \mu \rangle \le 0$ for $t > 0$ we have $\langle x^* - c, \mu \rangle = 0$. Thus,

$$\langle x - c, \mu \rangle \le 0 \text{ for all } x - c \in Z^+ \cap X$$

If we define the dual cone X_+^* by

$$X_+^* = \{x^* \in X^* : \langle x^*, x \rangle \le 0 \text{ for all } \quad x \in X \text{ and } x \le 0\},$$

then $\mu \in X_+^*$. Thus, the regular point condition gives the stronger regularity $\mu \in Z^+$ for the Lagrange multiplier μ.

<u>Example (Source Identification)</u> Let $X = Z = L^2(\Omega)$ and $S : \mathcal{L}(X, Y)$. and consider the constrained minimization:

$$\min \quad \frac{1}{2}(|Sf - y|_Y^2 + \alpha\, |f|^2)$$

subject to $-f \le 0$. In this case, $A = \alpha\, I + S^*S$. $\alpha > 0$, $a = S^*y$ and $G = -I$. The regular point condition (**??**) is given by

$$0 \in \operatorname{int}\{(-Z + f^*) - Z^- - f^*\},$$

Since $-Z + f^*$ contains Z^- and the regular point condition holds. Hence there exists $\mu \in Z^+$ such that

$$\alpha\, f^* + S^*Sf^* + \mu = S^*y, \quad \mu \ge 0 \text{ and } (-f^*, \mu) = 0.$$

For example, the operator S is defined for Poisson equation:

$$-\Delta u = f \text{ in } \Omega$$

with a boundary condition $\dfrac{\partial}{\partial u} + u = 0$ at $\Gamma = \partial\Omega$ and $Sf = Cu$ with a measurement operator C, e.g., $Cu = u|_\Gamma$ or $Cu = u|_{\hat{\Omega}}$ with $\hat{\Omega}$, a subdomain of Ω.

<u>Remark (Penalty Method)</u> Consider the penalized problem

$$J(x) + \frac{\beta}{2}|Ex - b|_Y^2 + \frac{\beta}{2}|\max(0, Gx - c)|_Z^2.$$

The necessary optimality condition is given by

$$Ax + E^*\lambda_\beta + G^*\mu_\beta = a, \quad \lambda_\beta = \beta\,(Ex - b), \quad \mu_\beta = \beta\,(Gx - c)^+.$$

It follows from Theorem 7 that there exists a unique solution $x_\beta$to (3.26) and satisfies

$$J(x_\beta) + \frac{\beta}{2}|Ex_\beta - b|_Y^2 + \frac{\beta}{2}|\max(0, Gx_\beta - c)|_Z^2 \leq J(x^*) \tag{3.35}$$

Thus, $\{x_\beta\}$ is bounded in X and has a weak convergence subsequence that converges to x in X. Moreover, we have

$$Ex = b \quad \text{and} \quad \max(0, Gx - c) = 0$$

and $J(x) \leq J(x^*)$ since the norm is weakly sequentially lower semi-continuous. Since the minimizer to (3.26) is unique, $x = x^*$ and $\lim_{\beta\to\infty} J(x_\beta) = J(x^*)$ which implies that $x_\beta \to x^*$ strongly in X.

Suppose $\lambda_\beta \in Y$ and $\mu_\beta \in Z$ are bounded, there exists a subsequence that converges weakly to (λ, μ) in $Y \times Z$ and $\mu \geq 0$. Since $\lim_{\beta\to\infty} J(x_\beta) = J(x^*)$ it follows from (3.35) that

$$(\mu_\beta, Gx_\beta - c) \to 0 \text{ as } \beta \to \infty$$

and thus $(\mu, Gx^* - c) = 0$. Hence (x^*, λ, μ) satisfies (3.34). Conversely, if there is a Lagrange multiplier $(\lambda^*, \mu^*) \in Y$, then $(\lambda_\beta, \mu_\beta)$ is bounded in $Y \times Z$. In fact, we have

$$(x_\beta - x^*, A(x_\beta - x^*)) + (x_\beta - x^*, E^*(\lambda_\beta - \lambda^*)) + (x_\beta - x^*, G^*(\mu_\beta - \mu^*)) = 0$$

where

$$\beta\,(x_\beta - x^*, E^*(\lambda_c - \lambda^*)) = \beta\,(Ex_\beta - b, \lambda_\beta - \lambda^*) = |\lambda_\beta|^2 - (\lambda_\beta, \lambda^*).$$

and

$$\beta\,(x_\beta - x^*, G^*(\mu_\beta - \mu^*)) = \beta\,(Gx_\beta - c - (Gx^* - c), \mu_\beta - \mu^*) \geq |\mu_\beta|^2 - (\mu_\beta, \mu^*).$$

Thus we obtain

$$\beta\,(x_\beta - x^*, A(x_\beta - x^*)) + |\lambda_\beta|^2 + |\mu_\beta|^2 \leq (\lambda_\beta, \lambda) + (\mu_\beta, \mu^*)$$

and hence $|(\lambda_\beta, \mu_\beta)| \leq |(\lambda^*, \mu^*)|$. It will be shown in Section 5 the regular point condition implies that $(\lambda_\beta, \mu_\beta)$ is uniformly bounded in $\beta > 0$.

Example (Quadratic Programming on R^n) Let $X = R^n$. Given $a \in R^n$, $b \in R^m$ and $c \in R^p$ consider

$$\min \quad \frac{1}{2}(Ax, x) - (a, x)$$

$$\text{subject to } Bx = b, \quad Cx \leq c$$

where $A \in R^{n\times n}$ is symmetric and positive, $B \in R^{m\times n}$ and $C \in R^{p\times n}$. Let C_a the (row) vectors of C corresponding to the active indices and assuming $\begin{pmatrix} B \\ C_a \end{pmatrix}$ is surjective, the necessary and sufficient optimality condition is given by

$$\begin{cases} Ax^* + B^t\lambda + C^t\mu = a \\ \\ Bx^* = b \\ \\ (Cx^* - c)_j \le 0 \quad \mu_j \ge 0 \quad \text{and} \quad (Cx^* - c)_j\mu_j = 0. \end{cases}$$

Example (Minimum norm problem with inequality constraint)

$$\min \quad |x|^2 \quad \text{subject to } (x, y_i) \le c_i,\ 1 \le \imath \le m.$$

Assume $\{y_i\}$ are linearly independent. Then it has a unique solution x^* and the necessary and sufficient optimality is given by

$$x^* + \sum_{i=1}^m \mu_i y_i = 0$$

$$(x^*, y_i) - c_i \le 0,\ \ \mu_i \ge 0 \text{ and } ((x^*, y_i) - c_i)\mu_i = 0,\ \ 1 \le i \le m.$$

That is, if G is Grammian ($G_{ij} = (y_i, y_j)$) then

$$(-Gx^* - c)_i \le 0 \ \ \mu_i \ge 0 \quad \text{and} \quad (-Gx^* - c)_i\mu_i = 0, \quad 1 \le i \le m.$$

Example (Pointwise Obstacle Problem) Consider

$$\min \quad \int_0^1 \frac{1}{2}|\frac{d}{dx}u(x)|^2 - \int_0^1 f(x)u(x)\,dx$$

$$\text{subject to } u(\frac{1}{2}) \le c$$

over $u \in X = H_0^1(0,1)$. Let $f = 1$. The Riesz representation of

$$F_1(u) = \int_0^1 f(t)u(t)\,dt, \quad F_2(u) = u(\frac{1}{2})$$

are

$$y_1(x) = \frac{1}{2}x(1-x). \quad y_2(x) = \frac{1}{2}(\frac{1}{2} - |x - \frac{1}{2}|),$$

respectively. Assume that $f = 1$. Then, we have

$$u^* - y_1 + \mu\, y_2 = 0, \quad u^*(\frac{1}{2}) = \frac{1}{8} - \frac{\mu}{4} \le c, \quad \mu \ge 0.$$

Thus, if $c \ge \frac{1}{8}$, then $u^* = y_1$ and $\mu = 0$. If $c \le \frac{1}{8}$, then $u^* = y_1 - \mu\, y_2$ and $\mu = 4(c - \frac{1}{8})$.

3.6 Constrained minimization in Banach spaces and Lagrange multiplier theory

In this section we discuss the general nonlinear programming. Let X, Y be Hilbert spaces and Z be a Hilbert lattice. Consider the constrained minimization;

$$\min \quad J(x) \quad \text{subject to } E(x) = 0 \text{ and } G(x) \le 0. \tag{3.36}$$

over a closed convex set $\mathcal{C}$ in X. Here, where $J : X \to R$, $E : X \to Y$ and $G : X \to Z$ are continuously differentiable.

Definition (Derivatives) (1) The functional F on a Banach space X is Gateaux differentiable at $u \in X$ if for all $d \in X$

$$\lim_{t \to 0} \frac{F(u + td) - F(u)}{t} = F'(u)(d)$$

exists. If the G-derivative $d \in X \to F'(u)(d) \in R$ is liner and bounded, then F is Frechet differentiable at u.
(2) $E : X \to Y$, Y ia Banach space is Frechet differentiable at u if there exists a linear bounded operator $E'(u)$ such that

$$\lim_{|d|_X \to 0} \frac{|E(u + d) - E(u) - E'(u)d|_Y}{|d|_X} = 0.$$

If Frechet derivative $u \in X \to E'(u) \in \mathcal{L}(X, Y)$ is continuous, then E is continuously differentiable.

Definition (Lower semi-continuous) (1) A functional F is lower-semi continuous if

$$\liminf_{n \to \infty} F(x_n) \geq F(\lim_{n \to \infty} x_n)$$

(2) A functional F is weakly lower-semi continuous if

$$\liminf_{n \to \infty} F(x_n) \geq F(w - \lim_{n \to \infty} x_n)$$

Theorem (Lower-semicontinuous) (1) Norm is weakly lower-semi continuous.
(2) A convex lower-semicontinuous functional is weakly lower-semi continuous.

Proof: Assume $x_n \to x$ weakly in X. Let $x^* \in F(x)$, i.e., $\langle x^*, x \rangle = |x^*||x|$. Then, we have

$$|x|^2 = \lim_{n \to \infty} \langle x^*, x_n \rangle$$

and

$$|\langle x^*, x_n \rangle| \leq |x_n||x^*|.$$

Thus,

$$\liminf_{n \to \infty} |x_n| \geq |x|.$$

(2) Since F is convex,

$$F(\sum_k t_k x_k) \leq \sum_k t_k F(x_k)$$

for all convex combination of x_k, i.e., $\sum t_k = 1,\ t_k \geq 0$. By the Mazur lemma there exists a sequence of convex combination of weak convergent sequence $(\{x_k\}, \{F(x_k)\})$ to $(x, F(x))$ in $X \times R$ that converges strongly to $(x, F(x))$ and thus

$$F(x) \leq \liminf n \to \infty \, F(x_n). \square$$

Theorem (Existence) (1) A lower-semicontinuous functional F on a compact set S has a minimizer.
(2) A weak lower-semicontinuous, coercive functional F has a minimizer.

Proof: (1) One can select a minimizing sequence $\{x_n\}$ such that $F(x_n) \downarrow \eta = \inf_{x \in S} F(x)$. Since S is a compact, there exits a subsequaence that converges strongly to $x^* \in S$. Since F is lower-semicontinuous, $F(x^*) \leq \eta$.
(2) Since $F(x) \to \infty$ as $|x| \to \infty$, there exits a bounded minimizing sequence $\{x_n\}$ of F. A bounded sequence in a reflexible Banach space X has a weak convergent subsequence to $x^* \in X$. Since F is weakly lower-semicontinuous, $F(x^*) \leq \eta$. $\square$

Next, we consider the constrained optimization in Banach spaces X, Y:

$$\min \quad F(y) \quad \text{subject to } G(y) \in K \tag{3.37}$$

over $y \in \mathcal{C}$, where K is a closed convex cone of a Banach space Y, $F : X \to R$ and $G : X \to Y$ are C_1 and $\mathcal{C}$ is a closed convex set of X. For example, if $K = \{0\}$, then $G(y) = 0$ is the equality constraint. If $Y = Z$ is a Hilbert lattice and

$$K = \{z \in Z : z \leq 0\},$$

then $G(y) \leq 0$ is the inequality constraint. Then, we have the Lagrange multiplier theory:

Theorem (Lagrange Multiplier Theory) Let $y^* \in \mathcal{C}$ be a solution to (3.37) and assume the regular point condition

$$0 \in \text{int}\,\{G'(y^*)(\mathcal{C} - y^*) - K + G(y^*)\} \tag{3.38}$$

holds. Then, there exists a Lagrange multiplier

$$\lambda \in K^+ = \{y \in Y : (y, z) \leq 0 \text{ for all } z \in K\} \text{ satisfying } (\lambda, G(y^*))_Y = 0$$

such that

$$(F'(y^*) + G'(y^*)^*\lambda, y - y^*)_X \geq 0 \quad \text{for all } y \in \mathcal{C}. \tag{3.39}$$

Proof: Let $\mathcal{F} = \{y \in \mathcal{C} : G(y) \in K\}$ be the feasible set. The sequential tangent cone of $\mathcal{F}$ at y^* is defined by

$$T(\mathcal{F}, y^*) = \{x \in X : x = \lim_{n\to\infty} \frac{y_n - y^*}{t_n},\ t_n \to 0^+,\ y_n \in \mathcal{F}\}.$$

It is easy show that

$$(F'(y^*), y - y^*) \geq 0 \quad \text{for all } y \in T(\mathcal{F}, y^*).$$

Let

$$\mathcal{C}(y^*) = \cup_{s \geq 0,\, y \in \mathcal{C}}\,\{s(y - y^*)\}, \quad K(G(y^*)) = \{K - t\,G(y^*),\ t \geq 0\}.$$

and define the linearizing cone $L(\mathcal{F}, y^*)$ of $\mathcal{F}$ at y^* by

$$L(\mathcal{F}, y^*) = \{z \in \mathcal{C}(y^*) : G'(y^*)z \in K(G(y^*))\}.$$

It follows the Listernik's theorem [] that

$$L(\mathcal{F}, y^*) \subseteq T(\mathcal{F}, y^*),$$

and thus

$$(F'(y^*), z) \geq 0 \quad \text{for all } z = y - y^* \in L(\mathcal{F}, y^*), \tag{3.40}$$

Define the convex cone B by

$$B = \{(F'(y^*)\mathcal{C}(y^*) + R^+, G'(y^*)\mathcal{C}(y^*) - K(G(y^*))\} \subset R \times Y.$$

By the regular point condition B has an interior point. From (3.40) the origin $(0,0)$ is a boundary point of B and thus there exists a hyperplane in $R \times Y$ that supports B at $(0,0)$, i.e., there exists a nontrivial $(\alpha, \lambda) \in R \times Y$ such that

$$\alpha \, (F'(y^*)z + r) + (\lambda, G'(y^*)z - y) \geq 0$$

for all $z \in \mathcal{C}(y^*)$, $y \in K(G(y^*))$ and $r \geq 0$. Setting $(z,r) = (0,0)$, we have $(\lambda, y) \leq 0$ for all $y \in K(G(y^*))$ and thus

$$\lambda \in K^+ \quad \text{and} \quad (\lambda, G(y^*)) = 0.$$

Letting $(r,y) = (0,0)$ and $z \in L(\mathcal{F}, y^*)$, we have $\alpha \geq 0$. If $\alpha = 0$, the regular point condition implies $\lambda = 0$, which is a contradiction. Without loss of generality we set $\alpha = 1$ and obtain (3.39) by setting $(r,y) = 0$. $\square$

Corollary (Nonlinear Programming) If there exists $\bar{x}$ such that $E(\bar{x}) = 0$ and $G(\bar{x}) \leq 0$, then there exits a solution x^* to (3.36). Assume the regular point condition

$$0 \in \text{int}\,\{E'(x^*)(\mathcal{C} - x^*) \times (G'(x^*)(\mathcal{C} - x^*) - Z^- + G(x^*))\},$$

then there exists $\lambda \in Y$ and $\mu \in Z$ such that

$$\begin{aligned}
&(J'(x^*) + (E'(x^*))^*\lambda + (G'(x^*)^*\mu, x - x^*) \geq 0 \quad \text{for all } x \in \mathcal{C}, \\
&E(x^*) = 0 \\
&G(x^*) \leq 0, \quad \mu \geq 0 \quad \text{and} \quad (G(x^*), \mu)_Z = 0.
\end{aligned} \tag{3.41}$$

Next, we consider the parameterized nonlinear programming. Let X, Y, U be Hilbert spaces and Z be a Hilbert lattice. We consider the constrained minimization of the form

$$\min \quad J(x,u) \quad \text{subject to } E(x,u) = 0 \text{ and } G(x,u) \leq 0 \text{ over } (x,u) \in X \times \hat{\mathcal{C}}, \tag{3.42}$$

where $J : X \times U \to R$, $E : X \times U \to Y$ and $G : X \times U \to Z$ are continuously differentiable and $\hat{\mathcal{C}}$ is a closed convex set in U. Here we assume that for given $u \in \hat{\mathcal{C}}$ there exists a unique $x = x(u) \in X$ that satisfies $E(x,u) = 0$. Thus, one can eliminate $x \in X$ as a function of $u \in U$ and thus (3.42) can be reduced to the constrained minimization over $u \in \hat{\mathcal{C}}$. But it is more advantageous to analyze (3.42) directly using Theorem (Lagrange Mulplier).

Corollary (Control Problem) Let (x^*, u^*) be a minimizer of (3.42) and assume the regular point holds. It follows from Theorem 8 that the necessary optimality condition is given by

$$\begin{cases}
J_x(x^*, u^*) + E_x(x^*, u^*)^*\lambda + G_x(x^*, u^*)^*\mu = 0 \\
(J_u(x^*, u^*) + E_u(x^*, u^*)^*\lambda + G_u(x^*, u^*)^*\mu, u - u^*) \geq 0 \text{ for all } u \in \hat{\mathcal{C}} \\
E(x^*, u^*) = 0 \\
G(x^*, \mu^*) \leq 0, \quad \mu \geq 0, \quad \text{and} \quad (G(x^*, u^*), \mu) = 0.
\end{cases} \tag{3.43}$$

Example (Control problem) Consider the optimal control problem

$$\begin{aligned}
&\min \quad \int_0^T (\ell(x(t)) + h(u(t)))\,dt + G(x(T)) \\
&\text{subject to } \frac{d}{dt}x(t) - f(x,u) = 0, \quad x(0) = x_0 \\
&u(t) \in \hat{U} = \text{a closed convex set in } R^m.
\end{aligned} \tag{3.44}$$

We let $X = H^1(0,T;R^n),\ U = L^2(0,T:R^m),\ Y = L^2(0,T;R^n)$ and

$$J(x,u) = \int_0^T (\ell(x) + h(u))\,dt + G(x(T)), \quad E = f(x,u) - \frac{d}{dt}x$$

and $\hat{\mathcal{C}} = \{u \in U,\ u(t) \in \hat{U} \text{ a.e in } (0,T)\}$. From (3.43) we have

$$\int_0^T (\ell'(x^*(t))\phi(t) + (p(t), f_x(x^*,u^*)\phi(t) - \frac{d}{dt}\phi)\,dt + G'(x(T))\phi(T) = 0$$

for all $\phi \in H^1(0,T;R^n)$ with $\phi(0) = 0$. By the integration by part

$$\int_0^T (\int_t^T (\ell'(x^*) + f_x(x^*,u^*)^t p)\,ds + G'(x^*(T)) - p(t), \frac{d\phi}{dt})\,dt = 0.$$

Since ϕ is arbitrary, we have

$$p(t) = G'(x^*(T)) + \int_t^T (\ell'(x^*) + f_x(x^*,u^*)^t p)\,ds$$

and thus the adjoint state p is absolutely continuous and satisfies

$$-\frac{d}{dt}p(t) = f_x(x^*,u^*)^t p(t) + \ell'(x^*), \quad p(T) = G'(x^*(T)).$$

From the the optimality, we have

$$\int_0^T (h'(u^*) + f_u(x^*,u^*)^t p(t))(u(t) - u^*(t))\,dt$$

for all $u \in \hat{\mathcal{C}}$, which implies

$$(h'(u^*(t)) + f_u(x^*(t),u^*(t))^t p(t))(v - u^*(t)) \geq 0 \text{ for all } v \in \hat{U}$$

at Lebesgue points t of $u^*(t)$. Hence we obtain the necessary optimality is of the form of Two-Point-Boundary value problem;

$$\begin{cases}
\frac{d}{dt}x^*(t) = f(x^*,u^*), \quad x(0) = x_0 \\
(h'(u^*(t)) + f_u(x^*(t),u^*(t))^t p(t), v - u^*(t)) \geq 0 \text{ for all } v \in \hat{U} \\
-\frac{d}{dt}p(t) = f_x(x^*,u^*)^t p(t) + l_x(x^*), \quad p(T) = G_x(x^*(T)).
\end{cases}$$

Let $\hat{U} = [-1, 1]$ coordinate-wise and $h(u) = \frac{\alpha}{2}|u|^2,\; f(x,u) = \tilde{f}(x) + Bu$. Thus, the optimality condition implies

$$(\alpha\, u^*(t) + B^t p(t), v - u^*(t))_{R^m} \geq 0 \text{ fur all } |v|_\infty \leq 1.$$

Or, equivalently $u^*(t)$ minimizes

$$\frac{\alpha}{2}\,|u|^2 + (B^t p(t), u) \text{ over } |u|_\infty \leq 1$$

Thus, we have

$$u^*(t) = \begin{cases} 1 & \text{if } -\frac{B^t p(t)}{\alpha} > 1 \\ \\ -\frac{B^t p(t)}{\alpha} & \text{if } |\frac{B^t p(t)}{\alpha}| \leq 1 \\ \\ -1 & \text{if } -\frac{B^t p(t)}{\alpha} < -1. \end{cases} \tag{3.45}$$

Example (Coefficient estimation) Let ω be a bounded open set in R^d and $\tilde{\Omega}$ be a subset in Ω. Consider the parameter estimation problem

$$\min \quad J(u,c) = \int_{\tilde{\Omega}} |y - u|\, dx^2 + \int_\Omega |\nabla c|^2\, dx$$

subject to

$$-\Delta u + c(x)u = f, \quad c \geq 0$$

over $(u,c) \in H_0^1(\Omega) \times H^1(\Omega)$. We let $Y = (H_0^1(\Omega))^*$ and

$$\langle E(u,c), \phi\rangle = (\nabla u, \nabla \phi) + (c\, u, \phi)_{L^2(\Omega)} - \langle f, \phi\rangle$$

for all $\phi \in H_0^1(\Omega)$. Since

$$\langle J_x(h) + E_x^* \lambda(h) = (\nabla h, \nabla \lambda) + (c\, h, \lambda)_{L^2(\omega)} - ((y - u)\, \chi_{\tilde{\Omega}}, h) = 0,$$

for all $h \in H_0^1(\Omega)$, i.e., the adjoint $\lambda \in H_0^1(\Omega)$ satisfies

$$-\Delta \lambda + c\, \lambda = (y - u)\, \chi_{\tilde{\Omega}}.$$

Also, there exits $\mu \geq 0$ in $L^2(\Omega)$ such that

$$\langle J_c + E_c^* \lambda, d\rangle = \alpha\, (\nabla c^*, \nabla d) + (d\, u, \lambda)_{L^2(\Omega)} - (\mu, d) = 0$$

for all $d \in H^1(\Omega)$ and

$$\langle \mu, c - c^*\rangle = 0 \text{ and } \langle \mu, d\rangle \geq 0 \text{ for all } d \in H^1(\Omega) \text{ satisfying } d \geq 0.$$

3.7 Minimum norm solution in Banach space

In this section we discuss the minimum distance and the minimum norm problem in Banach spaces. Given a subspace M of a normed space X, the orthogonal complement $M^\perp$ of M is defined by

$$M^\perp = \{x^* \in X^* : \langle x^*, x\rangle = 0 \text{ for all } x \in M\}.$$

Theorem If M is a closed subspace of a normed space, then

$$(M^\perp)^\perp = \{x \in X : \langle x, x^* \rangle = 0 \text{ for all } x^* \in M^\perp\} = M$$

Proof: If there exits a $x \in (M^\perp)^\perp$ but $x \notin M$, define a linear functional f by

$$f(\alpha\, x + m) = \alpha$$

on the space spanned by $x + M$. Since $x \notin m$ and M is closed

$$|f| = \sup_{m\in M,\, \alpha\in R} \frac{|f(\alpha\, x+m)|}{|\alpha\, x+m|_X} = \sup_{m\in M} \frac{|f(x+m)|}{|x+m|_X} = \frac{1}{\inf_{m\in M} |x+m|_X} < \infty,$$

and thus f is bounded. By the Hahn-Banach theorem, f can be extended to $x^* \in X^*$. Since $\langle x^*, m\rangle = 0$ for all $m \in M$, thus $x^* \in M^\perp$. It is contradicts to the fact that $\langle x^*, x\rangle = 1$. $\square$

Theorem (Duality I) Given $x \in X$ and a subspace M, we have

$$d = \inf_{m\in M} |x-m|_X = \max_{|x^*|\le 1,\, x^*\in M^\perp} \langle x^*, x\rangle = \langle x_0^*, x\rangle,$$

where the maximum is attained by some $x_0^* \in M^\perp$ with $|x_0^*| = 1$. If the infimum on the first equality holds for some $m_0 \in M$, then x_0^* is aligned with $x - m_0$.

Proof: By the definition of inf, for any $\epsilon > 0$, there exists $m_\epsilon \in M$ such that $|x - m_\epsilon| \le d + \epsilon$. For any $x^* \in M^\perp$, $|x^*| \le 1$, it follows that

$$\langle x^*, x\rangle = \langle x^*, x - m_\epsilon\rangle \le |x^*||x - m_\epsilon| \le d + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, $\langle x^*, x\rangle \le d$. It only remains to show that $\langle x_0^*, x\rangle = d$ for some $x_0^* \in X^*$. Define a linear functional f on the space spanned by $x + M$ by

$$f(\alpha\, x + m) = \alpha\, d$$

Since

$$|f| = \sup_{m\in M,\, \alpha\in R} \frac{|f(\alpha\, x+m)|}{|\alpha\, x+m|_X} = \frac{d}{\inf_{m\in M} |x+m|_X} = 1$$

f is bounded. By the Hahn-Banach theorem, there exists an extension x_0^* of f with $|x_0^*| = 1$. Since $\langle x_0^*, m\rangle = 0$ for all $m \in M$, $x_0^* \in M^\perp$. By construction $\langle x_0^*, x\rangle = d$. ???Moreover, if $|x - m_0| = d$ for some $m_0 \in M$, then

$$\langle x_0^*, x - m_0\rangle = d = |x_0^*||x - m_0|. \square$$

3.7.1 Duality Principle

Corollary (Duality) If $m_0 \in M$ attains the minimum if and only if there exists a nonzero $x^* \in M^\perp$ that is aligned with $x - m_0$.

Proof: Suppose that $x^* \in M^\perp$ is aligned with $x - m_0$,

$$|x - m_0| = \langle x^*, x - m_0\rangle = \langle x^*, x\rangle = \langle x^*, x\rangle \le |x - m|$$

for all $m \in M$. $\square$

Theorem (Duality principle) (1) Given $x^* \in X^*$, we have

$$\min_{m^* \in M^\perp} |x^* - m^*|_{X^*} = \sup_{x \in M,\ |x| \le 1} \langle x^*, x\rangle.$$

(2) The minimum of the left is achieved by some $m_0^* \in M^\perp$.
(3) If the supremum on the right is achieved by some $x_0 \in M$, then $x^* - m_0^*$ is aligned with x_0.

Proof: (1) Note that for $m^* \in\in M^\perp$

$$|x^* - m^*| = \sup_{|x| \le 1} \langle x^* - m^*, x\rangle \ge \sup_{|x| \le 1,\ x \in M} \langle x^* - m^*, x\rangle = \sup_{|x| \le 1,\ x \in M} \langle x^*, x\rangle.$$

Consider the restriction of x^* to M (with norm $\sup_{x \in M,\ |x| \le 1} \langle x^*, x\rangle$. Let y^* be the Hahn-Banach extension of this restricted x^*. Let $m_0^* = x^* - y^*$. Then $m_0^* \in M^\perp$ and $|x^* - m_0^*| = |y^*| = \sup_{x \in M,\ |x| \le 1} \langle x^*, x\rangle$. Thus, the minimum of the left is achieved by some $m_0^* \in M^\perp$. If the supremum on the right is achieved by some $x_0 \in M$, then $x^* - m_0^*$ is aligned with x_0. Obviously, $|x_0| = 1$ and $|x^* - m_0^*| = \langle x^*, x_0\rangle = \langle x^* - m_0^*, x_0\rangle$.
$\square$

In all applications of the duality theory, we use the alignment properties of the space and its dual to characterize optimum solutions, guarantee the existence of a solution by formulating minimum norm problems in the dual space, and examine to see if the dual problem is easier than the primal problem.
Example 1(Chebyshev Approximation) Problem is to to find a polynomial p of degree less than or equal to n that best approximates $f \in C[0,1]$ in the the sense of the sup-norm over $[a,b]$. Let M be the space of all polynomials of degree less than or equal to n. Then, M is a closed subspace of $C[0,1]$ of dimension $n+1$. The infimum of $|f-p|_\infty$ is achievable by some $p_0 \in M$ since M is closed. The Chebyshev theorem states that the set $\Gamma = \{t \in [0,1] : |f(t) - p_0(t)| = |f - p_0|_\infty\}$ contains at least $n+2$ points. In fact, $f - p_0$ must be aligned with some element in $M^\perp \subset C[0,1]^* = BV[0,1]$. Assume Γ contains $m < n+2$ points $0 \le t_1 < \cdots < t_m \le 1$. If $v \in BV[0,1]$ is aligned with $f - p_0$, then v is a piecewise continuous function with jump discontinuities only at these t_k's. Let t_k be a point of jump discontinuity of v. Define

$$q(t) = \Pi_{j \ne k}^m (t - t_j)$$

Then, $q \in M$ but $\langle q, v\rangle \ne 0$ and hence $v \notin M^\perp$, which is a contradiction.

Next, we consider the minimum norm problem;

$$\min \quad |x^*|_{X^*} \quad \text{subject to } \langle x^*, y_k\rangle = c_k,\ 1 \le k \le n.$$

where $y_1, \cdots, y_n \in X$ are given linearly independent vectors.

Theorem Let $\bar{x}^* \in X^*$ satisfying the constraints $\langle \bar{x}^*, y_k\rangle = c_k,\ 1 \le k \le n$. Let $M = span\{y_k\}$. Then,

$$d = \min_{\langle \bar{x}^*, y_k\rangle = c_k} |x^*|_{X^*} = \min_{m^* \in M^\perp} |\bar{x}^* - m^*|.$$

By the duality principle,

$$d = \min_{m^* \imath M^\perp} |\bar{x}^* - m^*| = \sup_{x \in M,\ |x| \le 1} \langle \bar{x}^*, x\rangle.$$

Write $x = Ya$ where $Y = [y_1, \cdots, y_n]$ and $a \in R^n$. Then

$$d = \min |x^*| = \max_{|Ya| \leq 1} \langle \bar{x}^*, Ya \rangle = \max_{|Ya| \leq 1} (c, a).$$

Thus, the optimal solution x_0^* must be aligned with Ya.

Example (DC motor control) Let $X = L^1(0,1)$ and $X^* = L^\infty(0,1)$.

$$\min \quad |u|_\infty$$

$$\text{subject to} \int_0^1 e^{t-1} u(t)\, dt = 0, \quad \int_0^1 (1 + e^{t-1}) u(t) = 1.$$

Let $y_1 = e^{t-1}$ and $y_2 = 1 + e^{t-1}$. From the duality principle

$$\min_{\langle y_k, u \rangle = c_k} |u|_\infty = \max_{|a_1 y_1 + a_2 y_2|_1 \leq 1} a_2.$$

The control problem now is reduced to finding constants a_1 and a_2 that maximizes a_2 subject to $\int_0^1 |(a_1 - a_2) e^{t-1} + a_2|\, dt \leq 1$. The optimal solution $u \in L^\infty(0,1)$ should be aligned with $(a_1 - a2) e^{t-1} + a_2 \in L^1(0,1)$. Since $(a_1 - a_2) e^{t-1} + a_2$, can change sign at most once. the alignment condition implies that u must have values $|u|_\infty$ and can change sign at most once, which is the so called bang-bang control.

4 Linear Operator Theory and C_0-semigroup

In this section we discuss the Cauchy problem

$$\frac{d}{dt} u(t) = Au(t) + f(t), \quad u(0) = u_0 \in X$$

in a Banach space X, where $u_0 \in X$ is the initial condition and $f \in L^1(0, T; X)$. We construct the mild solution $u(t) \in C(0, T; X)$:

$$u(t) = S(t) u_0 + \int_0^t S(t-s) f(s)\, ds \tag{4.1}$$

where a family of bounded linear operator $\{S(t),\ t \geq 0\}$ is C_0-semigroup on X.

Definition (C_0 semigroup) (1) Let X be a Banach space. A family of bounded linear operators $\{S(t),\ t \geq 0\}$ on X is called a strongly continuous (C_0) semigroup if

$$S(t+s) = S(t)S(s) \text{ for } t,\ s \geq 0 \text{ with } S(0) = I$$

$$|S(t)\phi - \phi| \to 0 \text{ as } t \to 0^+ \quad \text{for all } \phi \in X.$$

(2) A linear operator A in X defined by

$$A\phi = \lim_{t \to 0^+} \frac{S(t)\phi - \phi}{t} \tag{4.2}$$

with

$$dom(A) = \{\phi \in X : \text{the strong limit of } \lim_{t \to 0^+} \frac{S(t)\phi - \phi}{t} \text{ in } X \text{ exists}\}.$$

is called the infinitesimal generator of the C_0 semigroup $S(t)$.

In this section we present the basic theory of the linear C_0-semigroup on a Banach space X. The theory allows to analyze a wide class of the physical and engineering dynamics using the unified framework. We also present the concrete examples to demonstrate the theory. There is a necessary and sufficient condition (Hile-Yosida Theorem) for a closed, densely defined linear A in X to be the infinitesimal generator of the C_0 semigroup $S(t)$. Moreover, we will show that the mild solution $u(t)$ satisfies

$$\langle u(t), \psi\rangle = \langle u_0, \psi\rangle + \int (\langle x(s), A^*\psi\rangle + \langle f(s), \psi\rangle \, ds \tag{4.3}$$

for all $\psi \in dom\,(A^*)$.

Examples (1) For $A \in (X)$, define a sequence of linear operators in X

$$S_N(t) = \sum_k \frac{1}{k!}(At)^k.$$

Then

$$|S_N(t)| \le \sum \frac{1}{k!}(|A|t)^k \le e^{|A|\,t}$$

and

$$\frac{d}{S}_N(t) = AS_{N-1}(t)$$

Since $S(t) = e^{At} = \lim_{N\to\infty} S_N(t)$ in the operator norm, we have

$$\frac{d}{dt}S(t) = AS(t) = S(t)A.$$

(2) Consider the the hyperbolic equation

$$u_t + u_x = 0, \quad u(0,x) = u_0(x) \text{ in } (0,1). \tag{4.4}$$

Define the semigroup $S(t)$ of translations on $X = L^2(0,1)$ by

$$[S(t)u_0](x) = \tilde{u}_0(x-t), \quad \tilde{u}_0(x) = 0, x \le 0, \;\; \tilde{u}_0 = u_0 \text{ on } (0,1).$$

Then, $\{S(t),\ t \ge 0\}$ is a C_0 semigroup on X. If we define $u(t,x) = [S(t)u_0](x)$ with $u_0 \in H^1(0,1)$ with $u_0(0) = 0$ satisfies (4.5) a.e.. The generator A is given by

$$A\phi = -\phi' \text{ with } dom(A) = \{\phi \in H^1(0,1) \text{ with } u(0) = 0\}.$$

Thus, $u(t) = S(t)u_0$ satisfies the Cauchy problem $\frac{d}{u}(t) = Au(t)$ if $u_0 \in dom(A)$.
(3) Consider the the heat equation

$$u_t = \frac{\sigma^2}{2}\Delta u, \quad u(0,x) = u_0(x) \text{ in } L^2(R^n). \tag{4.5}$$

Define the semigroup

$$[S(t)u_0](x) = \frac{1}{(\sqrt{2\pi t}\sigma)^n}\int_{R^n} e^{-\frac{|x-y|^2}{2\sigma^2 t}}\, u_0(y)\, dy$$

Let A be closed, densely defined linear operator $dom(A) \to X$. We use the finite difference method in time to construct the mild solution (4.1). For a stepsize $\lambda > 0$ consider a sequence $\{u^n\}$ in X generated by

$$\frac{u^n - u^{n-1}}{\lambda} = Au^n + f^{n-1}, \tag{4.6}$$

with

$$f^{n-1} = \frac{1}{\lambda} \int_{(n-1)\lambda}^{n\lambda} f(t)\, dt.$$

Assume that for $\lambda > 0$ the resolvent operator

$$J_\lambda = (I - \lambda\, I)^{-1}$$

is bounded. Then, we have the product formula:

$$u^n = J_\lambda^n u_0 + \sum_{k=0}^{n-1} J_\lambda^{n-k} f^k\, \lambda. \tag{4.7}$$

In order to $u^n \in X$ is uniformly bounded in n for all $u_0 \in X$ (with $f = 0$), it is necessary that

$$|J_\lambda^n| \leq \frac{M}{(1-\lambda\omega)^n} \text{ for } \lambda\omega < 1, \tag{4.8}$$

for some $M \geq 1$ and $\omega \in R$.

Hile's Theorem Define a piecewise constant function in X by

$$u_\lambda(t) = u^{k-1} \text{ on } [t_{k-1}, t_k)$$

Then,

$$\max_{t\in[0,T]} |u_\lambda - u(t)|_x \to 0$$

as $\lambda \to 0^+$ to the mild solution (4.1). That is,

$$S(t)x = \lim_{n\to\infty} (I - \frac{t}{n} A)^{[\frac{t}{n}]} x$$

exists for all $x \in X$ and $\{S(t),\ t \geq 0\}$ is the C_0 semigurop on X and its generator is A.

Proof: First, note that

$$|J_\lambda| \leq \frac{M}{1-\lambda\omega}$$

and for $x \in dom(A)$

$$|J_\lambda x - x| = |\lambda J_\lambda A x| \leq \frac{\lambda}{1-\lambda\omega} |Ax| \to 0$$

as $\lambda \to 0^+$. Since $dom(A)$ is dense in X it follows that

$$|J_\lambda x - x| \to 0 \text{ as } \lambda \to 0^+ \text{ for all } x \in X.$$

Define the linear operators $T_\lambda(t)$ and $S_\lambda(t)$ by

$$S_\lambda(t) = J_\lambda^{k-1} \text{ and } T_\lambda(t) = J_\lambda^{k-1} + (t - t_k)(J_\lambda^k - J_\lambda^{k-1}), \text{ on } [t_{k-1}, t_k).$$

Then,

$$\frac{d}{dt} T_\lambda(t) = A S_\lambda(t), \text{ a.e } int \in [0, T].$$

Thus,

$$T_\lambda(t)u_0 - T_\mu(t)u_0 = \int_0^t \frac{d}{ds}(T_\lambda(s) T_\mu(t-s)u_0)\, ds = \int_0^t (S_\lambda(s)T_\mu(t-s) - T_\lambda(s)S_\lambda(t-s))Au_0\, ds$$

Since

$$T_\lambda(s)u - S_\lambda(s)u = (s - t_{k-1})(J_\lambda - I)T_\lambda(t_{k-1}) \text{ on } s \in [t_{k-1}, t_k).$$

By the bounded convergence theorem

$$|T_\lambda(t)u_0 - T_\mu(t)u_0|_X \to 0$$

as $\lambda,\ \mu \to 0$ for all $u_0 \in dom(A)$. Thus, the unique limit defines the linear operator $S(t)$ by

$$S(t)u_0 = \lim_{\lambda \to 0^+} S_\lambda(t)u_0. \tag{4.9}$$

Since

$$|S_\lambda(t)| \leq \frac{M}{(1 - \lambda\omega)^{[t/n]}} \leq M e^{\omega t},$$

where $[s]$ is the largest integer less than $s \in R$. and $dom(A)$ is dense, (4.9) holds for all $u_0 \in X$. Since

$$S(t+s)u = \lim_{\lambda \to 0^+} J_\lambda^{n+m} = J_\lambda^n J_\lambda^m u = S(t)S(s)u$$

and $\lim_{t\to 0^+} S(t)u = \lim_{t \to 0^+} J_t u = u$, $S(t)$ is the C_0 semigroup on X. Moreover, $\{S(t), t \geq 0\}$ is in the class $G(M, \omega)$, i.e.,

$$|S(t)| \leq M e^{\omega t}$$

Note that

$$T_\lambda(t)u_0 - u_0 = A \int_0^t S_\lambda u_0\, ds$$

Since $\lim_{\lambda \to 0^+} T_\lambda(t)u_0 = \lim_{\lambda\to 0^+} S_\lambda(t)u_0 = S(t)u_0$ and A is closed, we have

$$S(t)u_0 - u_0 = A \int_0^t S(s)u_0\, ds, \quad \int_0^t S(s)u_0\, ds \in dom(A).$$

If B is a generator of $\{S(t),\ t \geq 0\}$, then

$$Bx = \lim_{t \to 0^+} \frac{S(t)x - x}{t} = Ax$$

if $x \in dom(A)$. Conversely, if $u_0 \in dom(B)$, then $u_0 \in dom(A)$ since A is closed and $t \to S(t)u_0$ is continuous at 0 and thus

$$\frac{1}{t} \int_0^t S(s)u_0\, ds = u_0 \text{ as } t \to 0^+.$$

Hence
$$Ax = \frac{S(t)u_0 - u_0}{t} = Bx$$
That is, A is the generator of $\{S(t),\ t \geq 0\}$.

Similarly, we have
$$\sum_{k=0}^{n-1} J_\lambda^{n-k} f^k \ = \int_0^t S_\lambda(t-s)f(s)\,ds \to \int_0^t S(t-s)f(s)\,ds \text{ as } \lambda \to 0^+$$
by the Lebesque dominated convergence theorem. □

The following theorem state the basic properties of C_0 semigroups:

Theorem (Semigroup) (1) There exists $M \geq 1,\ \omega \in R$ such that $S \in G(M,\omega)$ class, i.e.,
$$|S(t)| \leq M\,e^{\omega t},\ \ t \geq 0. \tag{4.10}$$
(2) If $x(t) = S(t)x_0,\ x_0 \in X$, then $x \in C(0,T;X)$
(3) If $x_0 \in dom\,(A)$, then $x \in C^1(0,T;X) \cap C(0,T;dom(A))$ and
$$\frac{d}{dt}x(t) = Ax(t) = AS(t)x_0.$$
(4) The infinitesimal generator A is closed and densely defined. For $x \in X$
$$S(t)x - x = A\int_0^t S(s)x\,ds. \tag{4.11}$$
(5) $\lambda > \omega$ the resolvent is given by
$$(\lambda\,I - A)^{-1} = \int_0^\infty e^{-\lambda s}S(s)\,ds \tag{4.12}$$
with estimate
$$|(\lambda\,I - A)^{-n}| \leq \frac{M}{(\lambda-\omega)^n}. \tag{4.13}$$
Proof: (1) By the uniform boundedness principle there exists $M \geq 1$ such that $|S(t)| \leq M$ on $[0,t_0]$ For arbitrary $t = k\,t_0 + \tau$, $k \in N$ and $\tau \in [0,t_0)$ it follows from the semigroup property we get
$$|S(t)| \leq |S(\tau)||S(t_0|^k \leq Me^{k\,log\,|S(t_0)|} \leq Me^{\omega\,t}$$
with $\omega = \frac{1}{t_0}\,,log\,|S(t_0)|$.
(2) It follows from the semigroup property that for $h > 0$
$$x(t+h) - x(t) = (S(h) - I)S(t)x_0$$
and for $t - h \geq 0$
$$x(t-h) - x(t) = S(t-h)(I - S(h))x_0$$
Thus, $x \in C(0,T;X)$ follows from the strong continuity of $S(t)$ at $t = 0$.
(3)–(4) Moreover,
$$\frac{x(t+h)-x(t)}{h} = \frac{S(h)-I}{h}S(t)x_0 = S(t)\frac{S(h)x_0 - x_0}{h}$$

and thus $S(t)x_0 \in dom(A)$ and

$$\lim_{h\to 0^+} \frac{x(t+h)-x(t)}{h} = AS(t)x_0 = Ax(t).$$

Similarly,

$$\lim_{h\to 0^+} \frac{x(t-h)-x(t)}{-h} = \lim_{h\to 0^+} S(t-h)\frac{S(h)\phi-\phi}{h} = S(t)Ax_0.$$

Hence, for $x_0 \in dom(A)$

$$S(t)x_0 - x_0 = \int_0^t S(s)Ax_0\,ds = \int_0^t AS(s)x_0\,ds = A\int_0^t S(s)x_0\,ds \tag{4.14}$$

If $x_n \in don(A) \to x$ and $Ax_n \to y$ in X, we have

$$S(t)x - x = \int_0^t S(s)y\,ds$$

Since

$$\lim_{t\to 0^+} \frac{1}{t}\int_0^t S(s)y\,ds = y$$

$x \in dom(A)$ and $y = Ax$ and hence A is closed. Since A is closed it follows from (4.14) that for $x \in X$

$$\int_0^t S(s)x\,ds \in dom(A)$$

and (4.11) holds. For $x \in X$ let

$$x_h = \frac{1}{h}\int_0^h S(s)x\,ds \in dom(A)$$

Since $x_h \to x$ as $h \to 0^+$, $dom(A)$ is dense in X.
(5) For $\lambda > \omega$ define $R_t \in \mathcal{L}(X)$ by

$$R_t = \int_0^t e^{-\lambda s}S(s)\,ds.$$

Since $A - \lambda I$ is the infinitesimal generator of the semigroup $e^{\lambda t}S(t)$, from (4.11)

$$(\lambda I - A)R_t x = x - e^{-\lambda t}S(t)x.$$

Since A is closed and $|e^{-\lambda t}S(t)| \to 0$ as $t \to \infty$, we have $R = \lim_{t\to\infty} R_t$ satisfies

$$(\lambda I - A)R\phi = \phi.$$

Conversely, for $\phi \in dom(A)$

$$R(A-\lambda I)\phi = \int_0^\infty e^{-\lambda s}S(s)(A-\lambda I)\phi = \lim_{t\to\infty} e^{-\lambda t}S(t\phi - \phi = -\phi$$

Hence

$$R = \int_0^\infty e^{-\lambda s}S(s)\,ds = (\lambda I - A)^{-1}$$

Since for $\phi \in X$

$$|Rx| \le \int_0^\infty |e^{-\lambda s} S(s)x| \le M \int_0^\infty e^{(\omega-\lambda)s}|x|\,ds = \frac{M}{\lambda-\omega}|x|,$$

we have

$$|(\lambda\, I - A)^{-1}| \le \frac{M}{\lambda-\omega}, \quad \lambda > \omega.$$

Note that

$$(\lambda\, I - A)^{-2} = \int_0^\infty e^{-\lambda t} S(t)\,ds \int_0^\infty e^{\lambda\, s} S(s)\,ds = \int_0^\infty \int_0^\infty e^{-\lambda(t+s)} S(t+s)\,ds\,dt$$

$$= \int_0^\infty \int_t^\infty e^{-\lambda\sigma} S(\sigma)\,d\sigma\,dt = \int_0^\infty \sigma e^{-\lambda\sigma} S(\sigma)\,d\sigma.$$

By induction, we obtain

$$(\lambda\, I - A)^{-n} = \frac{1}{(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} S(t)\,dt. \tag{4.15}$$

Thus,

$$|(\lambda\, I - A)^{-n}| \le \frac{1}{(n-1)!} \int_0^\infty t^{n-1} e^{-(\lambda-\omega)t}\,dt = \frac{M}{(\lambda-\omega)^n}.\Box$$

We then we have the necessary and sufficient condition:

Hile-Yosida Theorem A closed, densely defined linear operator A on a Banach space X is the infinitesimal generator of a C_0 semigroup of class $G(M,\omega)$ if and only if

$$|(\lambda\, I - A)^{-n}| \le \frac{M}{(\lambda-\omega)^n} \quad \text{for all } \lambda > \omega \tag{4.16}$$

Proof: The sufficient part follows from the previous Theorem. But, we describe the Yosida construction. Define the Yosida approximation $A_\lambda \in \mathcal{L}(X)$ of A by

$$A_\lambda = \frac{J_\lambda - I}{\lambda} = AJ_\lambda. \tag{4.17}$$

Define the Yosida approximation:

$$S_\lambda(t) = e^{A_\lambda t} = e^{-\frac{t}{\lambda}} e^{J_\lambda \frac{t}{\lambda}}.$$

Since

$$|J_\lambda^k| \le \frac{M}{(1-\lambda\omega)^k}$$

we have

$$|S_\lambda(t)| \le e^{-\frac{t}{\lambda}} \sum_{k=0}^\infty \frac{M}{k!} |J_\lambda^k| (\frac{t}{\lambda})^k \le M e^{\frac{\omega}{1-\lambda\omega} t}.$$

Since

$$\frac{d}{ds} S_\lambda(s) S_{\hat{\lambda}}(t-s) = S_\lambda(s)(A_\lambda - A_{\hat{\lambda}}) S_{\hat{\lambda}}(t-s),$$

we have

$$S_\lambda(t)x - S_{\hat{\lambda}}(t)x = \int_0^t S_\lambda(s) S_{\hat{\lambda}}(t-s)(A_\lambda - A_{\hat{\lambda}})x\,ds$$

Thus, for $x \in dom(A)$

$$|S_\lambda(t)x - S_{\hat{\lambda}}(t)x| \leq M^2 t e^{\omega t} \, |(A_\lambda - A_{\hat{\lambda}})x| \to 0$$

as $\lambda, \ \hat{\lambda} \to 0^+$. Since $dom(A)$ is dense in X this implies that

$$S(t)x = \lim_{\lambda \to 0^+} S_\lambda(t)x \text{ exist for all } x \in X,$$

defines a C_0 semigroup of $G(M, \omega)$ class. The necessary part follows from (4.15) $\square$

Theorem (Mild solution) (1) If for $f \in L^1(0,T;X)$ define

$$x(t) = x(0) + \int_0^t S(t-s)f(s)\,ds,$$

then $x(t) \in C(0,T;X)$ and it satisfies

$$x(t) = A\int_0^t x(s)\,ds + \int_0^t f(s)\,ds. \tag{4.18}$$

(2) If $Af \in L^1(0,T;X)$ then $x \in C(0,T;dom(A))$ and

$$x(t) = x(0) + \int_0^t (Ax(s) + f(s))\,ds.$$

(3) If $f \in W^{1,1}(0,T;X)$, i.e. $f(t) = f(0) + \int_0^t f'(s)\,ds$, $\frac{d}{dt}f = f' \in L^1(0,T;X)$, then $Ax \in C(0,T;X)$ and

$$A\int_0^t S(t-s)f(s)\,ds = S(t)f(0) - f(t) + \int_0^t S(t-s)f'(s)\,ds. \tag{4.19}$$

Proof: Since

$$\int_0^t \int_0^\tau S(t-s)f(s)\,dsd\tau = \int_0^t (\int_s^t S(\tau - s)d\,\tau)f(s)\,ds,$$

and

$$A\int_0^t S(s)\,ds = S(t) - I$$

we have $x(t) \in dom(A)$ and

$$A\int_0^t x(s)\,ds = S(t)x - x + \int_0^t S(t-s)f(s)\,ds - \int_0^t f(s)\,ds.$$

and we have (4.18).
(2) Since for $h > 0$

$$\frac{x(t+h) - x(t)}{h} = \int_0^t S(t-s)\frac{S(h) - I}{h}f(s)\,ds + \frac{1}{h}\int_t^{t+h} S(t+h-s)f(s)\,ds$$

if $Af \in L^1(0,T;X)$

$$\lim_{h \to 0^+} \frac{x(t+h) - x(t)}{h} = \int_0^t S(t-s)Af(s)\,ds + f(t)$$

a.e. $t \in (0,T)$. Similarly,

$$\frac{x(t-h)-x(t)}{-h} = \int_0^{t-h} S(t-h-s)\frac{S(h)-I}{h} f(s)\,ds + \frac{1}{h}\int_{t-h}^t S(t-s)f(s)\,ds$$

$$\to \int_0^t S(t-s)Af(s)\,ds + f(t)$$

a.e. $t \in (0,T)$.
(3) Since

$$\frac{S(h)-I}{h}x(t) = \frac{1}{h}(\int_0^h S(t_h-s)f(s)\,ds - \int_t^{t+h} S(t+f-s)f(s)\,ds$$

$$+\int_0^t S(t-s)\frac{f(s+h)-f(s)}{h}\,ds,$$

letting $h \to 0^+$, we obtain(4.19). $\square$

It follows from Theorems the mild solution

$$x(t) = S(t)x(0) + \int_0^t S(t-s)f(s)\,ds$$

satisfies

$$x(t) = x(0) + A\int_0^t x(s) + \int_0^t f(s)\,ds.$$

Note that the mild solution $x \in C(0,T;X)$ depends continuously on $x(0) \in X$ and $f \in L^1(0,T;X)$ with estimate

$$|x(t)| \leq M(e^{\omega t}|x(0)| + \int_0^t e^{\omega(t-s)}|f(s)|\,ds).$$

Thus, the mild solution is the limit of a sequence $\{x_n\}$ of strong solutions with $x_n(0) \in dom(A)$ and $f_n \in W^{1,1}(0,T;X)$, i.e., since $dom(A)$ is dense in X and $W^{1,1}(0,T;X)$ is dense in $L^1(0,T;X)$,

$$|x_n(t)-x(t)|_X \to 0 \text{ uniformly on } [0,T]$$

for

$$|x_n(0)-x(0)|_x \to 0 \text{ and } |f_n - f|_{L^1(0,T;X)} \to 0 \text{ as } n \to \infty.$$

Moreover, the mild solution $x \in C(0,T:X)$ is a weak solution to the Cauchy problem

$$\frac{d}{dt}x(t) = Ax(t) + f(t) \tag{4.20}$$

in the sense of (4.3), i.e., for all $\psi \in dom(A^*)$ $\langle x(t),\psi\rangle_{X\times X^*}$ is absolutely continues and

$$\frac{d}{dt}\langle x(t),\psi\rangle = \langle x(t),\psi\rangle + \langle f(t),\psi\rangle \text{ a.e. in } (0,T).$$

If $x(0) \in dom(A)$ and $Af \in L^1(0,T;X)$, then $Ax \in C(0,T;X)$, $x \in W^{1,1}(0,T;X)$ and

$$\frac{d}{dt}x(t) = Ax(t) + f(t), \text{ a.e. in } (0,T)$$

If $x(0) \in dom(A)$ and $f \in W^{1,1}(0,T;X)$, then $x \in C(0,T;dom(A)) \cap C^1(0,T;X)$ and

$$\frac{d}{dt}x(t) = Ax(t) + f(t), \text{ everywhere in } [0,T].$$

The condition (4.16) is very difficult to check for a given A in general. For the case $M = 1$ we have a very complete characterization.

Lumer-Phillips Theorem The followings are equivalent:
(a) A is the infinitesimal generator of a C_0 semigroup of $G(1,\omega)$ class.
(b) $A - \omega\, I$ is a densely defined linear m-dissipative operator,i.e.

$$|(\lambda\, I - A)x| \geq (\lambda - \omega)|x| \quad \text{for all } x \in don(A), \;\; \lambda > \omega \tag{4.21}$$

and for some $\lambda_0 > \omega$

$$R(\lambda_0\, I - A) = X. \tag{4.22}$$

Proof: It follows from the m-dissipativity

$$|(\lambda_0\, I - A)^{-1}| \leq \frac{1}{\lambda_0 - \omega}$$

Suppose $x_n \in dom(A) \to x$ and $Ax_n \to y$ in X, the

$$x = \lim_{n\to\infty} x_n = (\lambda_0\, I - A)^{-1} \lim_{n\to\infty} (\lambda_0\, x_n - Ax_n) = (\lambda_0\, I - A)^{-1}(\lambda_0\, x - y).$$

Thus, $x \in dom(A)$ and $y = Ay$ and hence A is closed. Since for $\lambda > \omega$

$$\lambda\, I - A = (I + (\lambda - \lambda_0)(\lambda_0\, I - A)^{-1})(\lambda_0\, I - A),$$

if $\frac{|\lambda-\lambda_0|}{\lambda_0-\omega} < 1$, then $(\lambda\, I - A)^{-1} \in \mathcal{L}(X)$. Thus by the continuation method we have $(\lambda\, I - A)^{-1}$ exists and

$$|(\lambda\, I - A)^{-1}| \leq \frac{1}{\lambda - \omega}, \quad \lambda > \omega.$$

It follows from the Hile-Yosida theorem that (b) $\to$ (a).
(b) $\to$ (a) Since for $x^* \in F(x)$, the dual element of x, i.e. $x^* \in X^*$ satisfying $\langle x, x^*\rangle_{X\times X^*} = |x|^2$ and $|x^*| = |x|$

$$\langle e^{-\omega t}S(t)x, x^*\rangle \leq |x||x^*| = \langle x, x^*\rangle$$

we have for all $x \in dom(A)$

$$0 \geq \lim_{t\to 0^+} \langle \frac{e^{-\omega t}S(t)x - x}{t}, x^*\rangle = \langle (A - \omega\, I)x, x^*\rangle \text{ for all } x^* \in F(x).$$

which implies $A - \omega\, I$ is dissipative. $\Box$

Theorem (Dissipative I) (1) A is a ω-dissipative

$$|\lambda\, x - Ax| \geq (\lambda - \omega)|x| \text{ for all } x \in dom\,(A).$$

if and only if (2) for all $x \in dom(A)$ there exists $x^* \in F(x)$ such that

$$\langle Ax, x^*\rangle \leq \omega\, |x|^2. \tag{4.23}$$

(2) $\to$ (1). Let $x \in dom(A)$ and choose $x^* \in F(0)$ such that $\langle A, x^* \rangle \leq 0$. Then, for any $\lambda > 0$,

$$|x|^2 = \langle x, x^* \rangle = \langle \lambda\, x - Ax + Ax, x^* \rangle \leq \langle \lambda\, x - Ax, x^* \rangle - \omega\; |x|^2 \leq |\lambda\, x - Ax||x| - \omega\, |x|^2,$$

which implies (2).
(1) $\to$ (2). From (1) we obtain the estimate

$$-\frac{1}{\lambda}(|x - \lambda\, Ax| - |x|) \leq 0$$

and

$$\langle Ax, x \rangle_- = \lim_{\lambda \to 0^+} = \frac{1}{\lambda}(|x| - |x - \lambda\, Ax|) \leq 0$$

which implies there exists $x^* \in F(x)$ such that (4.23) holds since $\langle Ax, x \rangle_- = \langle AX, x^* \rangle$ for some $x^* \in F(x)$. $\square$

Thus, Lumer-Phillips theorem says that if m-diisipative, then (4.23) hold for all $x^* \in F(x)$.

Theorem (Dissipative II) Let A be a closed densely defined operator on X. If A and A^* are dissipative, then A is m-dissipative and thus the infinitesimal generator of a $C-0$-semigroup of contractions.

Proof: Let $y \in \overline{R(I-A)}$ be given. Then there exists a sequence $x_n \in dom(A)$ such that $y_= x_n - Ax_n \to y$ as $n \to \infty$. By the dissipativity of A we obtain

$$|x_n - x_m| \leq |x_n - x_m - A(x_n - x_m)| \leq |y_- y_m|$$

Hence x_n is a Cauchy sequence in X. We set $x = lim_{n \to \infty} x_n$. Since A is closed, we see that $x \in dom\,(A)$ and $x - Ax = y$, i.e., $y \in R(I-A)$. Thus $R(I-A)$ is closed. Assume that $R(I-A) \neq X$. Then there exists an $x^* \in X^*$ such that

$$\langle (I-A)x, x^*) = 0 \text{ for all } x \in dom\,(A).$$

By definition of the dual operator this implies $x^* \in dom\,(A^*)$ and $(IA)^* x^* = 0$. Dissipativity of A* implies $|x^*| < |x^* - A^* x^*| = 0$, which is a contradiction. $\square$

Example (revisited example 1)

$$A\phi = -\phi' \text{ in } X = L^2(0,1)$$

and for $\phi \in H^1(0,1)$

$$(A\phi, \phi)_X = -\int_0^1 \phi'(x)\phi\, dx == \frac{1}{2}(|\phi(0)|^2 - |\phi(1)|^2)$$

Thus, A is dissipative if and only if $\phi(0) = 0$, the in flow condition. Define the domain of A by

$$dom(A) = \{\phi \in H^1(0,1) : \phi(0) = 0\}$$

The resolvent equation is equivalent to

$$\lambda\, u + \frac{d}{dx} u = f$$

and

$$u(x) = \int_0^x e^{-\lambda\,(x-s)} f(s)\,ds,$$

and $R(\lambda\,I - A) = X$. By the Lumer-Philips theorem A generates the C_0 semigroup on $X = L^2(0,1)$.

Example (Conduction equation) Consider the heat conduction equation:

$$\frac{d}{dt}u = Au = \sum_{i,j} a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + b(x)\cdot\nabla u + c(x)u, \quad \text{in } \Omega.$$

Let $X = C(\Omega)$ and $dom(A) \subset C^2(\Omega)$. Assume that $a \in R^{n\times n} \in C(Omega)$ $b \in R^{n,1}$ and $c \in R$ are continuous on $\bar{\Omega}$ and a is symmetric and

$$m\,I \le a(x) \le M\,I \text{ for } 0 < m \le M < \infty.$$

Then, if x_0 is an interior point of Ω at which the maximum of $\phi \in C^2(\Omega)$ is attained. Then,

$$\nabla\phi(x_0) = 0, \;\; \sum_{ij} a_{ij}(x_0)\frac{\partial^2 u}{\partial x_i \partial x_j}(x_0) \le 0.$$

and thus

$$(\lambda\,\phi - A\phi)(x_0) \le \omega\,\phi(x_0)$$

where

$$\omega \le \max_{x\in\Omega} c(x).$$

Similarly, if x_0 is an interior point of Ω at which the minimum of $\phi \in C^2(\Omega)$ is attained, then

$$(\lambda\,\phi - A\phi)(x_0) \ge 0$$

If $x_0 \in \partial\Omega$ attains the maximum, then

$$\frac{\partial}{\partial\nu}\phi(x_0) \le 0.$$

Consider the domain with the Robin boundary condition:

$$dom(A) = \{u \in \alpha(x)\,u(x) + \beta(x)\frac{\partial}{\partial\nu}u = 0 \text{ at } \partial\Omega\}$$

with $\alpha,\ \beta \ge 0$ and $\inf_{x\in\partial\Omega}(\alpha(x) + \beta(x)) > 0$. Then,

$$|\lambda\phi - A\phi|_X \ge (\lambda - \omega)|\phi|_X.$$

for all $\phi \in dom(A)$, which shows A is dissipative. The range condition follows from the Lax Milgram theory.

Example (Advection equation) Consider the advection equation

$$u_t + (\vec{b}(x)u) = \nu\,\Delta u.$$

Let $X = L^1(\Omega)$. Assume

$$\vec{b} \in L^\infty(\Omega)$$

Let $\rho \in C^1(R)$ be a monotonically increasing function satisfying $\rho(0) = 0$ and $\rho(x) = sign(x),\ |x| \geq 1$ and $\rho_\epsilon(s) = \rho(\frac{s}{\epsilon})$ for $\epsilon > 0$. For $u \in C^1(\Omega)$

$$(Au, u) = \int_\Gamma (\nu \frac{\partial}{\partial n} u - n \cdot \vec{b}\, u, \rho_\epsilon(u))\, ds + (\vec{b}\, u - \nu\, u_x, \frac{1}{\epsilon}\rho'_\epsilon(u)\, u_x) + (c\, u, \rho_\epsilon(u)).$$

where

$$(\vec{b}\, u, \frac{1}{\epsilon}\rho'_\epsilon(u)\, u_x) \leq \nu\, (u_x, \frac{1}{\epsilon}\rho'_\epsilon(u)\, u_x) + \frac{\epsilon}{4\nu}\, meas(\{|u| \leq \epsilon\})$$

Assume the inflow condition u on $\{s \in \partial\Omega : n \cdot b < 0\}$. Note that for $u \in L^1(R^d)$

$$(u, \rho_\epsilon(u)) \to |u|_1 \quad \text{and} \quad (\psi, \rho_\epsilon(u)) \to (\psi, sign_0(u)) \ \text{ for } \ \psi \in L^1(\Omega)$$

as $\epsilon \to 0$. It thus follows

$$(\lambda - \omega)\, |u| \leq |\lambda\, u - \lambda\, Au|. \tag{4.24}$$

Since $C^2(\Omega)$ is dense in $W^{2,1}(R^d)$ (4.24) holds for $u \in W^{2,1}(\Omega)$.

Example ($X = L^p(\Omega)$(Let $Au = \nu\, \Delta u + b \cdot \nabla u$ with homogeneous boundary condition $u = 0$ on $X = L^p(\Omega)$. Since

$$\langle \Delta u, u^* \rangle = \int_\Omega \Delta, |u|^{p-2} u) = -(p-1) \int_\Omega (\nabla u, |u|^{p-2} \nabla u)$$

and

$$(b \cdot \nabla u, |u|^{p-2} u)_{L^2} \leq \frac{(p-1)\nu}{2} |(\nabla u, |u|^{p-2} \nabla u)_{L^2} + \frac{|b|_\infty^2}{2\nu(p-1)}\, (|u|^p, 1)_{L^2}$$

we have

$$\langle Au, u^* \rangle \omega |u|^2$$

for some $\omega > 0$.

Example (Second order equation) Let $V \subset H = H^* \subset V^*$ be the Gelfand triple. Let ρ be a bounded bilinear form on $H \times H$, μ and σ be bounded bilinear forms on $V \times V$. Assume ρ and σ are symmetric and coercive and $\mu(\phi, \phi) \geq 0$ for all $\phi \in V$. Consider the second order equation

$$\rho(u_{tt}, \phi) + \mu(u_t, \phi) + \sigma(u, \phi) = \langle f(t), \phi \rangle \text{ for all } \phi \in V.$$

Define linear operators M (mass), D (dampping), K and (stiffness) by

$$(M\phi, \psi)_H = \rho(\phi, \psi), \quad \phi,\ \psi \in H \langle D\phi, \psi \rangle = \mu(\phi, \psi) \quad \phi,\ \psi \in V \langle K\phi, \psi \rangle_{V^* \times V}, \quad \phi,\ \psi \in V$$

Let $v = u_t$ and define A on $X = V \times H$ by

$$A(u, v) = (v, -M^{-1}(Ku + Dv))$$

with domain

$$dom\, (A) = \{(u, v) \in X : v \in V \text{ and } Ku + Dv \in H\}$$

The state space X is a Hilbert space with inner product

$$((u_1, v_1), (u_, v_2)) = \sigma(u_1, u_2) + \rho(v_1, v_2)$$

and

$$E(t) = |(u(t), v(t))|_X^2 = \sigma(u(t), u(t)) + \rho(v(t), v(t))$$

defines the energy of the state $x(t) = (u(t), v(t))$. First, we show that A is dissipative:

$$(A(u,v),(u,v))_X = \sigma(u,v)+\rho(-M^{-1}(Ku+Dv),v) = \sigma(u,v)-\sigma(u,v)-\mu(v,v) = -\mu(v,v) \leq 0$$

Next, we show that $R(\lambda I - A) = X$. That is, for $(f,g) \in X$ the exists a solution $(u,v) \in dom\,(A)$ satisfying

$$\lambda\, u - v = f, \quad \lambda M v + Dv + Ku = Mg,$$

or equivalently $v = \lambda\, u - f$ and

$$\lambda^2 Mu + \lambda Du + Ku = Mg + \lambda\, Mf + Df \tag{4.25}$$

Define the bilinear form a on $V \times V$

$$a(\phi,\psi) = \lambda^2\, \rho(\phi,\psi) + \lambda\, \mu(\phi,\psi) + \sigma(\phi,\psi)$$

Then, a is bounded and coercive and if we let

$$F(\phi) = (M(g + \lambda\, f)\phi)_H + \mu(f,\phi)$$

then $F \in V^*$. It thus follows from the Lax-Milgram theory there exists a unique solution $u \in V$ to (4.25) and $Dv + Ku \in H$.

For example, consider the wave equation

$$\frac{1}{c^2(x)} u_{tt} = \Delta u, \quad \frac{\partial u}{\partial n} + \kappa(x) u_t = 0 \text{ and } \partial\Omega.$$

In this example we let $V = H^1(\Omega)/R$ and $H = L^2(\Omega)$ and define

$$\sigma(\phi,\psi) = \int_\Omega (\nabla\phi, \nabla\psi)\, dx$$

$$\mu(\phi,\psi) = \int_{\partial\Omega} \kappa(x)\phi, \psi\, ds$$

$$\rho(\phi,\psi) = \int_\Omega \frac{1}{c^2(x)}\, \phi\psi\, dx.$$

Example (Maxwell equation)

$$\epsilon\, E_t = \nabla \times H$$

$$\mu\, H_t = -\nabla \times E$$

with boundary condition

$$E \times n = 0$$

The dissipativity follows from

$$E \cdot (\nabla \times H) + H \cdot (\nabla \times E) = div(E$$

Theorem (Dual semigroup) Let X be a reflexive Banach space. The adjoint $S^*(t)$ of the C_0 semigroup $S(t)$ on X forms the C_0 semigroup and the infinitesimal generator of $S^*(t)$ is A^*. Let X be a Hilbert space and $dom(A^*)$ be the Hilbert space with graph norm and X_{-1} be the strong dual space of $dom(A^*)$, then the extension $S(t)$ to X_{-1} defines the C_0 semigroup on X_{-1}.

Proof: (1) Since for $t,\ s \geq 0$

$$S^*(t+s) = (S(s)S(t))^* = S^*(t)S^*(s)$$

and

$$\langle x, S^*(t)y - y\rangle_{X\times X^*} = \langle S(t)x - x, y\rangle_{X\times X^*} \to 0.$$

for $x \in X$ and $y \in X^*$ Thus, $S^*(t)$ is weakly star continuous at $t = 0$ and let B is the generator of $S^*(t)$ as

$$Bx = w^* - \lim \frac{S^*(t)x - x}{t}.$$

Since

$$(\frac{S(t)x - x}{t}, y) = (x, \frac{S^*(t)y - y}{t}),$$

for all $x \in dom(A)$ and $y \in dom(B)$ we have

$$\langle Ax, y\rangle_{X\times X^*} = \langle x, By\rangle_{X\times X^*}$$

and thus $B = A^*$. Thus, A^* is the generator of a w^*- continuous semigroup on X^*.
(2) Since

$$S^*(t)y - y = A^* \int_0^t S^*(s)y\,ds$$

for all $y \in Y = \overline{dom\,(A^*)}$. Thus, $S^*(t)$ is strongly continuous at $t = 0$ on Y.
(3) If X is reflexive, $\overline{dom\,(A^*)} = X^*$. If not, there exists a nonzero $y_0 \in X$ such that $\langle y_0, x^*\rangle_{X\times X^*} = 0$ for all $x^* \in dom(A^*)$. Thus, for $x_0 = (\lambda\, I - A)^{-1} y_0$ $\langle \lambda\, x_0 - Ax_0, x^*\rangle = \langle x_0, \lambda\, x^* - A^*x^*\rangle = 0$. Letting $x^* = (\lambda\, I - A^*)^{-1} x_0^*$ for $x_0^* \in F(x_0)$, we have $x_0 = 0$ and thus $y_0 = 0$, which yields a contradiction.
(4) $X_1 = dom\,(A^*)$ is a closed subspace of X^* and is a invariant set of $S^*(t)$. Since A^* is closed, $S^*(t)$ is the C_0 semigroup on X_1 equipped with its graph norm. Thus,

$$(S^*(t))^* \text{ is the } C_0 \text{ semigroup on } X_{-1} = X_1^*$$

and defines the extension of $S(t)$ to X_{-1}. Since for $x \in X \subset X_{-1}$ and $x^* \in X^*$

$$\langle S(t)x, x^*\rangle = \langle x, S^*(t)x^*\rangle,$$

$S(t)$ is the restriction of $(S^*(t))^*$ onto X. $\square$

4.1 Sectorial operator and Analytic semigroup

In this section we have the representation of the semigroup $S(t)$ in terms of the inverse Laplace transform. Taking the Laplace transform of

$$\frac{d}{dt}x(t) = Ax(t) + f(t)$$

we have

$$\hat{x} = (\lambda\, I - A)^{-1}(x(0) + \hat{f})$$

where for $\lambda > \omega$

$$\hat{x} = \int_0^\infty e^{-\lambda t} x(t)\, dt$$

is the Laplace transform of $x(t)$. We have the following the representation theory (inverse formula).

Theorem (Resolvent Calculus) For $x \in dom(A^2)$ and $\gamma > \omega$

$$S(t)x = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\lambda t} (\lambda\, I - A)^{-1} x\, d\lambda. \tag{4.26}$$

Proof: Let A_μ be the Yosida approximation of A. Since $\mathrm{Re}\,\sigma(A_\mu) \le \dfrac{\omega_0}{1 - \mu\omega_0} < \gamma$, we have

$$u_\mu(t) = S_\mu(t)x = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{\lambda t} (\lambda\, I - A_\mu)^{-1} x\, d\lambda.$$

Note that

$$\lambda(\lambda\, I - A)^{-1} = I + (\lambda\, I - A)^{-1} A. \tag{4.27}$$

Since

$$\frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} \frac{e^{\lambda t}}{\lambda}\, d\lambda = 1$$

and

$$\int_{\gamma - i\infty}^{\gamma + i\infty} |(|\lambda - \omega|^{-2}|\, d\lambda < \infty,$$

we have

$$|S_\mu(t)x| \le M\, |A^2 x|,$$

uniformly in $\mu > 0$. Since

$$(\lambda\, I - A_\mu)^{-1}x - (\lambda\, I - A)^{-1}x = \frac{\mu}{1 + \lambda\mu} (\nu\, I - A)^{-1} (\lambda\, I - A)^{-1} A^2 x,$$

where $\nu = \dfrac{\lambda}{1 + \lambda\mu}$, $\{u_\mu(t)\}$ is Cauchy in $C(0, T; X)$ if $x \in dom(A^2)$. Letting $\mu \to 0^+$, we obtain (4.26). $\square$

Next we consider the sectorial operator. For $\delta > 0$ let

$$\Sigma_\omega^\delta = \{\lambda \in C : arg(\lambda - \omega) < \frac{\pi}{2} + \delta\}$$

be the sector in the complex plane C. A closed, densely defined, linear operator A on a Banach space X is a sectorial operator if

$$|(\lambda\, I - A)^{-1}| \le \frac{M}{|\lambda - \omega|} \text{ for all } \lambda \in \Sigma_\omega^\delta.$$

For $0 < \theta < \delta$ let $\Gamma = \Gamma_{\omega,\theta}$ be the integration path defined by

$$\Gamma^{\pm} = \{z \in C : |z| \geq \delta,\ arg(z-\omega) = \pm(\tfrac{\pi}{2}+\theta)\},$$

$$\Gamma_0 = \{z \in C : |z| = \delta,\ |arg(z-\omega)| \leq \tfrac{\pi}{2}+\theta\}.$$

For $0 < \theta < \delta$ define a family $\{S(t),\ t \geq 0\}$ of bounded linear operators on X by

$$S(t)x = \frac{1}{2\pi i}\int_\Gamma e^{\lambda t}(\lambda\, I - A)^{-1}x\, d\lambda. \tag{4.28}$$

Theorem (Analytic semigroup) If A is a sectorial operator on a Banach space X, then A generates an analytic (C_0) semigroup on X, i.e., for $x \in X$ $t \to S(t)x$ is an analytic function on $(0,\infty)$. We have the representation (4.28) for $x \in X$ and

$$|AS(t)x|_X \leq \frac{M_\theta}{t}|x|_X$$

Proof:

The elliptic operator A defined by the Lax-Milgram theorem defines a sectorial operator on Hilbert space.

Theorem (Sectorial operator) Let $V,\ H$ are Hilbert spaces and assume $H \subset V^*$. Let $\rho(u,v)$ is bounded bilinear form on $H \times H$ and

$$\rho(u,u) \geq |u|_H^2 \text{ for all } u \in H$$

Let $a(u,v)$ to be a bounded bilinear form on $V \times V$ with

$$\sigma(u,u) \geq \delta\, |u|_V^2 \text{ for all } u \in V$$

Define the linear operator A by

$$\rho(Au,\phi) = a(u,\phi) \text{ for all } \phi \in V.$$

Then, for $Re\,\lambda > 0$we have

$$|(\lambda\, I - A)^{-1}|_{\mathcal{L}(V,V^*)} \leq \tfrac{1}{\delta}$$

$$|(\lambda\, I - A)^{-1}|_{\mathcal{L}(H)} \leq \tfrac{M}{|\lambda|}$$

$$|(\lambda\, I - A)^{-1}|_{\mathcal{L}(V^*,H)} \leq \tfrac{M}{\sqrt{|\lambda|}}$$

$$|(\lambda\, I - A)^{-1}|_{\mathcal{L}(H,V)} \leq \tfrac{M}{\sqrt{|\lambda|}}$$

Proof: Let $a(u,v)$ to be a bounded bilinear form on $V \times V$. For $f \in V^*$ and $\Re\lambda > 0$, Define $M\mathcal{H},\mathcal{H}$ by

$$(Mu,v) = \rho(u,v) \text{ for all } v \in H$$

and $A_0 \in \mathcal{L}(V,V^*)$ by

$$\langle A_0 u, v\rangle = \sigma(u,v) \text{ for } v \in V$$

Then, $(\lambda I - A)u = M^{-1}f$ is equivalent to

$$\lambda\rho(u,\phi) + a(u,\phi) = \langle f,\phi\rangle,\ \text{for all } \phi \in V. \tag{4.29}$$

It follows from the Lax-Milgram theorem that (4.29) has a unique solution, given $f \in V^*$ and

$$Re\,\lambda\rho(u,u) + a(u,u) \le |f|_{V^*}|u|_V.$$

Thus,

$$|(\lambda I - A)^{-1}|_{\mathcal{L}(V^*,V)} \le \frac{1}{\delta}.$$

Also,

$$|\lambda|\,|u|_H^2 \le |f|_{V^*}|u|_V + M\,|u|_V^2 = M_1\,|f|_{V^*}^2$$

for $M_1 = 1 + \frac{M}{\delta^2}$ and thus

$$|(\lambda I - A)^{-1}|_{\mathcal{L}(V^*,H)} \le \frac{\sqrt{M_1}}{|\lambda|^{1/2}}.$$

For $f \in H \subset V^*$

$$\delta\,|u|_V^2 \le Re\,\lambda\,\rho(u,u) + a(u,u) \le |f|_H|u|_H, \tag{4.30}$$

and

$$|\lambda|\rho(u,u) \le |f|_H|u|_H + M|u|_V^2 \le M_1|f|_H|u|_H$$

Thus,

$$|(\lambda I - A)^{-1}|_{\mathcal{L}(H)} \le \frac{M_1}{|\lambda|}.$$

Also, from (4.30)

$$\delta\,|u|_V^2 \le |f|_H|u|_H \le M_1\,|f|^2.$$

which implies

$$|(\lambda I - A)^{-1}|_{\mathcal{L}(H,V)} \le \frac{M_2}{|\lambda|^{1/2}}.$$

4.2 Approximation Theory

In this section we discuss the approximation theory for the C_0-semigroup.

Theorem (Trotter-Kato theorem) Let X and X_n be Banach spaces and A and A_n be the infinitesimal generator of C_0 semigroups $S(t)$ on X and $S_n(t)$ on X_n of $G(M,\omega)$ class. Assume a family of uniformly bounded linear operators $P_n \in \mathcal{L}(X,X_n)$ and $E_n \in \mathcal{L}(X_n,X)$ satisfy

$$P_nE_n = I \quad |E_nP_nx - x|_X \to 0 \text{ for all } x \in X$$

Then, the followings are equivalent.
(1) there exist a $\lambda_0 > \omega$ such that for all $x \in X$

$$|E_n(\lambda_0\,I - A_n)^{-1}P_nx - (\lambda_0\,I - A)^{-1}x|_X \to 0 \quad \text{as } n \to \infty,$$

(2) For every $x \in X$ and $T \ge 0$

$$|E_nS_n(t)P_nx - S(t)x|_X \to \quad \text{as } n \to \infty.$$

Proof: Since for $\lambda > \omega$

$$E_n(\lambda\, I - A)^{-1}P_n x - (\lambda\, I - A)^{-1}x = \int_0^\infty E_n S_n(t)P_n x - S(t)x\, dt$$

(1) follows from (2). Conversely, from the representation theory

$$E_n S_n(t)P_n x - S(t)x = \frac{1}{2\pi i}\int_{\gamma - i\infty}^{\gamma + i\infty}(E_n(\lambda\, I - A)^{-1}P_n x - (\lambda\, I - A)^{-1}x)\, d\lambda$$

where

$$(\lambda\, I - A)^{-1} - (\lambda_0\, I - A)^{-1} = (\lambda - \lambda_0)(\lambda\, I - A)^{-1}(\lambda_0\, I - A)^{-1}.$$

Thus, from the proof of Theorem (Resolvent Calculus) (1) holds for $x \in dom(A^2)$. But since $dom(A^2)$ is dense in X, (2) implies (1). $\square$

<u>Example (Trotter-Kato theoarem)</u> Consider the heat equation on $\Omega = (0,1)\times(0,1)$:

$$\frac{d}{dt}u(t) = \Delta u, \quad u(0,x) = u_0(x)$$

with boundary condition $u = 0$ at the boundary $\partial\Omega$. We use the central difference approximation on uniform grid points: $(i\, h, j\, h) \in \Omega$ with mesh size $h = \frac{1}{n}$:

$$\frac{d}{dt}u_{i,j}(t) = \Delta_h u = \frac{1}{h}\left(\frac{u_{i+1,j} - u_{i,j}}{h} - \frac{u_{i,j} - u_{i-1,j}}{h}\right) + \frac{1}{h}\left(\frac{u_{i,j+1} - u_{i,j}}{h} - \frac{u_{i,j} - u_{i,j-1}}{h}\right)$$

for $1 \le i,\ j \le n_1$, where $u_{i,0} = u_{i,n} = u1, j = u_{n,j} = 0$ at the boundary node. First, let $X = C(\Omega)$ and $X_n = R^{(n-1)^2}$ with sup norm. Let $E_n u_{i,j}$ = the piecewise linear interpolation and $(P_n u)_{i,j} = u(i\, h, j\, h)$ is the pointwise evaluation. We will prove that Δ_h is dissipative on X_n. Suppose $u_{ij} = |u_n|_\infty$. Then, since

$$\lambda\, u_{i,j} - (\Delta_h u)_{i,j} = f_{ij}$$

and

$$-(\Delta_h u)_{i,j} = \frac{1}{h^2}(4u_{i,j} - u_{i+1,j} - u_{i,j+1} - u_{i-1,j} - u_{i,j-1}) \ge 0$$

we have

$$0 \le u_{i,j} \le \frac{f_{i,j}}{\lambda}.$$

Thus, Δ_h is dissipative on X_n with sup norm.

Next $X = L^2(\Omega)$ and X_n with ℓ^2 norm. Then,

$$(-\Delta_h u_n, u_n) = \sum_{i,j} \left|\frac{u_{i,j} - u_{i-1,j}}{h}\right|^2 + \left|\frac{u_{i,j} - u_{i,j-1}}{h}\right|^2 \ge 0$$

and thus Δ_h is dissipative on X_n with ℓ^2 norm.

<u>Example</u>

5 Monotone Operator Theory and Nonlinear semigroup

In this section we discuss the nonlinear operator theory that extends the Lax-Milgram theory for nonlinear monotone operators and mappings from Banach space from X to X^*. Also, we extend the C_0 semigroup theory to the nonlinear case.

Let us denote by $F : X \to X^*$, the duality mapping of X, i.e.,

$$F(x) = \{x^* \in X^* : \langle x, x^* \rangle = |x|^2 = |x^*|^2\}.$$

By Hahn-Banach theorem, $F(x)$ is non-empty. In general F is multi-valued. Therefore, when X is a Hilbert space, $\langle \cdot, \cdot \rangle$ coincides with its inner product if X^* is identified with X and $F(x) = x$.

Let H be a Hilbert space with scalar product (ϕ, ψ) and X be a real, reflexive Banach space and $X \subset H$ with continuous dense injection. Let X^* denote the strong dual space of X. H is identified with its dual so that $X \subset H = H^* \subset X^*$. The dual product $\langle \phi, \psi \rangle$ on $X \times X^*$ is the continuous extension of the scalar product of H restricted to $X \times H$.

The following proposition contains some further important properties of the duality mapping F.

Theorem (Duality Mapping) (a) $F(x)$ is a closed convex subset.
(b) If X^* is strictly convex (i.e., balls in X^* are strictly convex), then for any $x \in X$, $F(x)$ is single-valued. Moreover, the mapping $x \to F(x)$ is demicontinuous, i.e., if $x_n \to x$ in X, then $F(x_n)$ converges weakly star to $F(x)$ in X^*.
(c) Assume X be uniformly convex (i.e., for each $0 < \epsilon < 2$ there exists $\delta = \delta(\epsilon) > 0$ such that if $|x| = |y| = 1$and $|x - y| > \epsilon$, then $|x + y| \le 2(1 - \delta)$). If x_n converges weakly to x and $\limsup_{n\to\infty} |x_n| \le |x|$, then x_n converges strongly to x in X.
(d) If X^* is uniformly convex, then the mapping $x \to F(x)$ is uniformly continuous on bounded subsets of X.

Proof: (a) Closeness of $F(x)$ is an easy consequence of the follows from the continuity of the duality product. Choose $x_1^*, x_2^* \in F(x)$ and $\alpha \in (0, 1)$. For arbitrary $z \in X$ we have (using $|x_1^*| = |x_2^*| = |x|$) $\langle z, \alpha x_1^* + (1 - \alpha)x_2^* \rangle \le \alpha|z|\,|x_1^*| + (1 - \alpha)|z|\,|x_2^*| = |z|\,|x|$, which shows $|\alpha x_1^* + (1 - \alpha)x_2^*| \le |x|$. Using $\langle x, x^* \rangle = \langle x, x_1^* \rangle = |x|^2$ we get $\langle x, \alpha x_1^* + (1 - \alpha)x_2^* \rangle = \alpha\langle x, x_1^* \rangle + (1 - \alpha)\langle x, x_2^* \rangle = |x|^2$, so that $|\alpha x_1^* + (1 - \alpha)x_2^*| = |x|$. This proves $\alpha x_1^* + (1 - \alpha)x_2^* \in F(x)$.
(b) Choose $x_1^*,\ x_2^* \in F(x)$, $\alpha \in (0, 1)$ and assume that $|\alpha x_1^* + (1-\alpha)x_2^*| = |x|$. Since X^* is strictly convex, this implies $x_1^* = x_2^*$. Let $\{x_n\}$ be a sequence such that $x_n \to x \in X$. From $|F(x_n)| = |x_n|$ and the fact that closed balls in X^* are weakly star compact we see that there exists a weakly star accumulation point x^* of $\{F(x_n)\}$. Since the closed ball in X^* is weakly star closed, thus

$$\langle x, x^* \rangle = |x|^2 \ge |x^*|^2.$$

Hence $\langle x, x^* \rangle = |x|^2 = |x^*|^2$ and thus $x^* = F(x)$. Since $F(x)$ is single-valued, this implies $F(x_n)$ converges weakly to $F(x)$.
(c) Since $\liminf |x_n| \le |x|$, thus $\lim_{n\to\infty} |x_n| = |x|$. We set $y_n = x_n/|x_n|$ and $y = x/|x|$. Then y_n converges weakly to y in X. Suppose y_n does not converge strongly to y in

X. Then there exists an $\epsilon > 0$ such that for a subsequence $y_{\tilde{n}}$ $|y_{\tilde{n}} - y| > \epsilon$. Since X^* is uniformly convex there exists a $\delta > 0$ such that $|y_{\tilde{n}} + y| \leq 2(1-\delta)$. Since the norm is weakly lower semicontinuos, letting $\tilde{n} \to \infty$ we obtain $|y| \leq 1-\delta$, which is a contradiction.
(d) Assume F is not uniformly continuous on bounded subsets of X. Then there exist constants $M > 0$, $\epsilon > 0$ and sequences $\{u_n\}, \{v_n\}$ in X satisfying

$$|u_n|, |v_n| \leq M, \quad |u_n - v_n| \to 0, \quad \text{and} \quad |F(u_n) - F(v_n)| \geq \epsilon.$$

Without loss of the generality we can assume that, for a constant $\beta > 0$, we have in addition $|u_n| \geq \beta$, $|v_n| \geq \beta$. We set $x_n = u_n/|u_n|$ and $y_n = v_n/|v_n|$. Then we have

$$|x_n - y_n| = \frac{1}{|u_n|\,|v_n|}\,||v_n|u_n - |u_n|v_n|$$

$$\leq \frac{1}{\beta^2}\left(|v_n||u_n - v_n| + ||v_n| - |u_n||\,|v_n|\right) \leq \frac{2M}{\beta^2}|u_n - v_n| \to 0 \quad \text{as } n \to \infty.$$

Obviously we have $2 \geq |F(x_n) + F(y_n)| \geq \langle x_n, F(x_n) + F(y_n)\rangle$ and this together with

$$\langle x_n, F(x_n) + F(y_n)\rangle = |x_n|^2 + |y_n|^2 + \langle x_n - y_n, F(y_n)\rangle$$

$$= 2 + \langle x_n - y_n, F(y_n)\rangle \geq 2 - |x_n - y_n|$$

implies

$$\lim_{n\to\infty} |F(x_n) + F(y_n)| = 2.$$

Suppose there exists an $\epsilon_0 > 0$ and a subsequence $\{n_k\}$ such that $|F(x_{n_k}) - F(y_{n_k})| \geq \epsilon_0$. Observing $|F(x_{n_k})| = |F(y_{n_k})| = 1$ and using uniform convexity of X^* we conclude that there exists a $\delta_0 > 0$ such that

$$|F(x_{n_k}) + F(y_{n_k})| \leq 2(1-\delta_0),$$

which is a contradiction to the above. Therefore we have $\lim_{n\to\infty} |F(x_n) - F(y_n)| = 0$. Thus

$$|F(u_n) - F(v_n)| \leq |u_n|\,|F(x_n) - F(y_n)| + ||u_n| - |v_n||\,|F(y_n)|$$

which implies $F(u_n)$ converges strongly to $F(v_n)$. This contradiction proves the result. □

Definition (Monotone Mapping)
(a) A mapping $A \subset X \times X^*$ be given. is called *monotone* if

$$\langle x_1 - x_2, y_1 - y_2\rangle \geq 0 \quad \text{for all } [x_1, y_1], [x_2, y_2] \in A.$$

(b) A monotone mapping A is called *maximal monotone* if any monotone extension of A coincides with A, i.e., if for $[x, y] \in X \times X^*$, $\langle x - u, y - v\rangle \geq 0$ for all $[u, v] \in A$ then $[x, y] \in A$.
(c) The operator A is called *coercive* if for all sequences $[x_n, y_n] \in A$ with $\lim_{n\to\infty} |x_n| = \infty$ we have

$$\lim_{n\to\infty} \frac{\langle x_n, y_n\rangle}{|x_n|} = \infty.$$

(d) Assume that A is single-valued with $\mathrm{dom}(A) = X$. The operator A is called *hemicontinuous*on X if for all $x_1, x_2, x \in X$, the function defined by

$$t \in R \to \langle x, A(x_1 + t x_2)\rangle$$

is continuous on R.

For example, let F be the duality mapping of X. Then F is monotone, coercive and hemicontinuous. Indeed, for $[x_1, y_1], [x_2, y_2] \in F$ we have

$$\langle x_1 - x_2, y_1 - y_2\rangle = |x_1|^2 - \langle x_1, y_2\rangle - \langle x_2, y_1\rangle + |x_2|^2 \geq (|x_1| - |x_2|)^2 \geq 0, \tag{5.1}$$

which shows monotonicity of F. Coercivity is obvious and hemicontinuity follows from the continuity of the duality product.

<u>Lemma 1</u> Let X be a finite dimensional Banach space and A be a hemicontinuous monotone operator from X to X^*. Then A is continuous.

Proof: We first show that A is bounded on bounded subsets. In fact, otherwise there exists a sequence $\{x_n\}$ in X such that $|Ax_n| \to \infty$ and $x_n \to x_0$ as $n \to \infty$. By monotonicity we have

$$\langle x_n - x, \frac{Ax_n}{|Ax_n|} - \frac{Ax}{|Ax_n|}\rangle \geq 0 \quad \text{for all } \; x \in X.$$

Without loss of generality we can assume that $\frac{Ax_n}{|Ax_n|} \to y_0$ in X^* as $n \to \infty$. Thus

$$\langle x_0 - x, y_0\rangle \geq 0 \quad \text{for all } \; x \in X$$

and therefore $y_0 = 0$. This is a contradiction and thus A is bounded. Now, assume $\{x_n\}$ converges to x_0 and let y_0 be a cluster point of $\{Ax_n\}$. Again by monotonicity of A

$$\langle x_0 - x, y_0 - Ax\rangle \geq 0 \quad \text{for all } \; x \in X.$$

Setting $x = x_0 + t\,(u - x_0),\; t > 0$ for arbitrary $u \in X$, we have

$$\langle x_0 - u, y_0 - A(x_0 + t\,(u - x_0)) \geq 0\rangle \quad \text{for all } \; u \in X.$$

Then, letting limit $t \to 0^+$, by hemicontinuity of A we have

$$\langle x_0 - u, y_0 - Ax_0\rangle \geq 0 \quad \text{for all } \; u \in X,$$

which implies $y_0 = Ax_0$. $\square$

<u>Lemma 2</u> Let X be a reflexive Banach space and $A : X \to X^*$ be a hemicontinuous monotone operator. Then A is maximumal monotone.

Proof: For $[x_0, y_0] \in X \times X^*$

$$\langle x_0 - u, y_0 - Au\rangle \geq 0 \quad \text{for all } \; u \in X.$$

Setting $u = x_0 + t\,(x - x_0),\; t > 0$ and letting $t \to 0^+$, by hemicontinuity of A we have

$$\langle x_0 - x, y_0 - Ax_0\rangle \geq 0 \quad \text{for all } \; x \in X.$$

Hence $y_0 = Ax_0$ and thus A is maximum monotone. $\square$

The next theorem characterizes maximal monotone operators by a range condition.

<u>**Minty–Browder Theorem**</u> Assume that X, X^* are reflexive and strictly convex. Let F denote the duality mapping of X and assume that $A \subset X \times X^*$ is monotone. Then A is maximal monotone if and only if

$$\mathit{Range}\,(\lambda F + A) = X^*$$

for all $\lambda > 0$ or, equivalently, for some $\lambda > 0$.

Proof: Assume that the range condition is satisfied for some $\lambda > 0$ and let $[x_0, y_0] \in X \times X^*$ be such that

$$\langle x_0 - u, y_0 - v \rangle \geq 0 \quad \text{for all } [u, v] \in A.$$

Then there exists an element $[x_1, y_1] \in A$ with

$$\lambda F x_1 + y_1 = \lambda F x_0 + y_0. \tag{5.2}$$

From these we obtain, setting $[u, v] = [x_1, y_1]$,

$$\langle x_1 - x_0, F x_1 - F x_0 \rangle \leq 0.$$

By monotonicity of F we also have the converse inequality, so that

$$\langle x_1 - x_0, F x_1 - F x_0 \rangle = 0.$$

From (5.10) this implies that $|x_1| = |x_0|$ and $\langle x_1, F x_0 \rangle = |x_1|^2$, $\langle x_0, F x_1 \rangle = |x_0|^2$. Hence $F x_0 = F x_1$ and

$$\langle x_1, F x_0 \rangle = \langle x_0, F x_0 \rangle = |x_0|^2 = |F x_0|^2.$$

If we denote by F^* the duality mapping of X^* (which is also single-valued), then the last equation implies $x_1 = x_0 = F^*(F x_0)$. This and (5.11) imply that $[x_0, y_0] = [x_1, y_1] \in A$, which proves that A is maximal monotone. $\square$

In stead of the detailed proof of ”only if’ part of Theorem, we state the following results. $\square$

<u>**Corollary**</u> Let X be reflexive and A be a monotone, everywhere defined, hemicontinous operator. If A is coercive, then $R(A) = X^*$.

Proof: Suppose A is coercive. Let $y_0 \in X^*$ be arbitrary. By the Appland's renorming theorem, we may assume that X and X^* are strictly convex Banach spaces. It then follows from Theorem that every $\lambda > 0$, equation

$$\lambda\, F x_\lambda + A x_\lambda = y_0$$

has a solution $x_\lambda \in X$. Multiplying this by x_λ,

$$\lambda\, |x_\lambda|^2 + \langle x_\lambda, A x_\lambda \rangle = \langle y_0, x_\lambda \rangle.$$

and thus

$$\frac{\langle x_\lambda, Ax_\lambda \rangle}{|x_\lambda|_X} \leq |y_0|_{X^*}$$

Since A is coercive, this implies that $\{x_\lambda\}$ is bounded in X as $\lambda \to 0^+$. Thus, we may assume that x_λ converges weakly to x_0 in X and Ax_λ converges strongly to y_0 in X^* as $\lambda \to 0^+$. Since A is monotone

$$\langle x_\lambda - x, y_0 - \lambda F x_\lambda - Ax \rangle \geq 0,$$

and letting $\lambda \to 0^+$, we have

$$\langle x_0 - x, y_0 - Ax \rangle \geq 0,$$

for all $x \in X$. Since A is maximal monotone, this implies $y_0 = Ax_0$. Hence, we conclude $R(A) = X^*$. $\square$

Theorem (Galerkin Approximation) Assume X is a reflexive, separable Banach space and A is a bounded, hemicontinuous, coercive monotone operator from X into X^*. Let $X_n = span\{\phi\}_{i=1}^n$ satisfies the density condition: for each $\psi \in X$ and any $\epsilon > 0$ there exists a sequence $\psi_n \in X_n$ such that $|\psi - \psi_n| \to 0$ as $n \to \infty$. The x_n be the solution to

$$\langle \psi, Ax_n \rangle = \langle \psi, f \rangle \quad \text{for all } \ \psi \in X_n, \tag{5.3}$$

then there exists a subsequence of $\{x_n\}$ that converges weakly to a solution to $Ax = f$.

Proof: Since $\langle x, Ax \rangle / |x|_X \to \infty$ as $|x|_X \to \infty$ there exists a solution x_n to (5.14) and $|x_n|_X$ is bounded. Since A is bounded, thus Ax_n bounded. Thus there exists a subsequence of $\{n\}$ (denoted by the same) such that x_n converges weakly to x in X and Ax_n converges weakly in X^*. Since

$$\lim_{n \to \infty} \langle \psi, Ax_n \rangle = \lim_{n \to \infty} (\langle \psi_n, f \rangle + \langle \psi - \psi_n, Ax_n \rangle) = \langle \psi, f \rangle$$

Ax_n converges weakly to f. Since A is monotone

$$\langle x_n - u, Ax_n - Au \rangle \geq 0 \quad \text{for all } \ u \in X$$

Note that

$$\lim_{n \to \infty} \langle x_n, Ax_n \rangle = \lim_{n \to \infty} \langle x_n, f \rangle = \langle x, f \rangle.$$

Thus taking limit $n \to \infty$, we obtain

$$\langle x - u, f - Au \rangle \geq 0 \text{ for all } \ u \in X.$$

Since A is maximum monotone this implies $Ax = f$. $\square$

The main theorem for monotone operators applies directly to the model problem involving the p-Laplace operator

$$-div(|\nabla u|^{p-2} \nabla u) = f \text{ on } \Omega$$

(with appropriate boundary conditions) and

$$-\Delta u + c\,u = f, \quad -\frac{\partial}{\partial n} u \in \beta(u). \text{ at } \partial\Omega$$

with β maximal monotone on R. Also, nonlinear problems of non-variational form are applicable, e.g.,

$$Lu + F(u, \nabla u) = f \text{ on } \Omega$$

where

$$L(u) = -div(\sigma(\nabla u) - \vec{b}\, u)$$

and we are looking for a solution $u \in W_0^{1,p}(\Omega),\ 1 < p < \infty$. We assume the following conditions:
(i) Monotonicity for the principle part $L(u)$:

$$(\sigma(\xi) - \sigma(\eta), \xi - \eta)_{R^n} \geq 0 \text{ for all } \xi,\ \eta \in R^n.$$

(ii) Monotonicity for $F = F(u)$:

$$(F(u) - F(v), u - v) \geq 0 \text{ for all } u,\ v \in R.$$

(iii) Coerciveness and Growth condition: for some $c,\ d > 0$

$$(\sigma(\xi), \sigma) \geq c\, |\xi|^p, \quad |\sigma(\xi)| \leq d\, (1 + |\xi|^{p-1})$$

hold for all $\xi \in R^n$.

5.1 Pseudo-Monotone operator

Next, we examine a somewhat more general class of nonlinear operators, called pseudo-monotone operators. In applications, it often occurs that the hypotheses imposed on monotonicity are unnecessarily strong. In particular, the monotonicity assumption $F(u)$ involves both the first-order derivatives and the function itself. In general the compactness will take care of the lower-order terms.

Definition (Pseudo-Monotone Operator) Let X be a reflexive Banach space. An operator $T : X \to X^*$ is said to be pseudo-monotone operator if T is a bounded operator (not necessarily continuous) and if whenever u_n converges weakly to u in X and

$$\limsup_{n\to\infty} \langle u_n - u, T(u_n)\rangle \leq 0,$$

it follows that, for all $v \in X$,

$$\liminf_{n\to\infty} \langle u_n - v, T(u_n)\rangle \geq \langle u - v, T(u)\rangle. \tag{5.4}$$

For example, a monotone and hemi-continuous operator is pseudo-monotone.

Lemma If T is maximal monotone, then T is pseudo-monotone.

Proof: Since T is monotone,

$$\langle u_n - u, T(u_n)\rangle \geq \langle u_n - u, T(u)\rangle \to 0$$

it follows from (5.4) that $\lim_n \langle u_n - u, T(u_n)\rangle = 0$. Assume $u_n \to u$ and $T(u_n) \to y$ weakly in X and X^*, respectively. Thus,

$$0 \leq \langle u_n - v, T(u_n) - T(v)\rangle = \langle u_n - u + u - v, T(u_n) - T(v)\rangle \to \langle u - v, y - T(v)\rangle.$$

Since T is maximal monotone, we have $y = T(u)$. Now,

$$\lim_n \langle u_n - v, T(u_n)\rangle = \lim_n \langle u_n - u + u - v, T(u_n)\rangle = \langle u - v, T(u)\rangle. \square$$

A class of pseudo-monotone operators is intermediate between monotone and hemicontinuous. The sum of two pseudo-monotone operators is pseudo-monotone. Moreover, the sum of a pseudo-monotone and a strongly continuous operator (x_n converges weakly to x, then $Tx_n \to Tx$ strongly) is pseudo-monotone.

Property of Pseudo-monotone operator
(1) Letting $v = u$ in (5.4)

$$0 \leq liminf \langle u_n - u, T(u_n)\rangle \leq \limsup \langle u_n - u, T(u_n)\rangle \leq 0$$

and thus $\lim_n \langle u_n - u, T(u_n)\rangle = 0$.
(2) From (1) and (5.4)

$$\langle u - v, Tu\rangle \leq \liminf \langle u_n - v, T(u_n)\rangle + \liminf \langle u - u_n - u, T(u_n)\rangle = \liminf \langle u - v, T(u_n)\rangle$$

(3) $T(u_n) \to T(u)$ weakly. In fact, from(2)

$$\langle v, T(u)\rangle \leq \liminf \langle v, T(u_n)\rangle \leq \limsup \langle v, T(u_n)\rangle$$

$$= \limsup \langle u_n - u, T(u_n)\rangle + \lim sup \langle u + v - u_n, T(u_n)\rangle$$

$$= -\liminf \langle u_n - (u + v), T(u_n)\rangle \leq \langle v, T(u)\rangle.$$

(4) $\langle u_n - u, T(u_n) - T(u)\rangle \to 0$ By the hypothesis

$$0 \leq \liminf \langle u_n - u, T(u_n)\rangle = \liminf \langle u_n - u, T(u_n)\rangle - \liminf \langle u_n - u, T(u)\rangle$$

$$= \liminf \langle u_n - u, T(u_n) - T(u)\rangle \leq \limsup \langle u_n - u, T(u_n) - T(u)\rangle$$

$$= \limsup \langle u_n - u, T(u_n) - T(u)\rangle \leq 0$$

(5) $\langle u_n, T(u_n)\rangle \to \langle u, T(u)\rangle$. Since

$$\langle u_n, T(u_n)\rangle = \langle u_n - u, T(u_n) - u\rangle - \langle u, T(u_n)\rangle - \langle u_n, T(u)\rangle + \langle u_n, T(u_n)\rangle$$

the claim follows from (3)–(4).
(6) Conversely, if $T(u_n) \to T(u)$ weakly and $\langle u_n, T(u_n)\rangle \to \langle u, T(u)\rangle \to 0$, then the hypothesis holds.

Using a very similar proof to that of the Browder-Minty theorem, one can show the following.

Theorem (Bresiz) Assume, the operator $A : X \to X^*$ is pseudo-monotone, bounded and coercive on the real separable and reflexive Banach space X. Then, for each $f \in X^*$ the equation $A(u) = f$ has a solution.

5.2 Generalized Pseudo-Monotone Mapping

In this section we discuss pseudo-monotone mappings. We extend the notion of the maximal monotone mapping, i.e., the maximal monotone mapping is a generalized pseudo-monotone mappings.

Definition (Pseudo-monotone Mapping)
(a) The set $T(u)$ is nonempty, bounded, closed and convex for all $u \in X$.
(b) T is upper semicontinuous from each finite-dimensional subspace F of X to the weak topology on X^*, i.e., to a given element $u+ \in F$ and a weak neighborhood V of $T(u)$, in X^* there exists a neighborhood U of u_0 in F such that $T(u) \subset V$ for all $u \in U$.
(c) If $\{u_n\}$ is a sequence in X that converges weakly to$u \in X$ and if $w_n \in T(u_n)$ is such that $\limsup\langle u_n - u, w_n\rangle \leq 0$, then to each element $v \in X$ there exists $w \in T(u)$ with the property that

$$\liminf\langle u_n - v, w_n\rangle \geq \langle u - v, w\rangle.$$

Definition (Generalized Pseudo-monotone Mapping) A mapping T from X into X^* is said to be generalized pseudo-monotone if the following is satisfied. For any sequence $\{u_n\}$ in X and a corresponding sequence $\{w_n\}$ in X^* with $w_n \in T(x_n)$, if u_n converges weakly to u and w_n weakly to w such that

$$\limsup_{n\to\infty}\langle u_n - u, w_n)\rangle \leq 0,$$

then $w \in T(u)$ and $\langle u_n, w_n\rangle \to \langle u, w\rangle$.

Let X be a reflexive Banach space, T a pseudo-monotone mapping from X into X^*. Then T is generalized pseudo-monotone. Conversely, suppose T is a bounded generalized pseudo-monotone mapping from X into X^*. and assume that for each $u \in X$, Tu is a nonempty closed convex subset of X^*. Then T is pseudo-monotone.
PROPOSITION 2. A maximal monotone mapping T from the Banach space X into X^* is generalized pseudo-monotone.

Proof: Let $\{u_n\}$ be a sequence in X that converges weakly to $u \in X$, $\{w_n\}$ a sequence in X^* with $w_n \in T(u_n)$ that converges weakly to $w \in X^*$. Suppose that $\limsup\langle u_n - u, w_n\rangle \leq 0$ Let $[x, y] \in T$ be an arbitrary element of the graph $G(T)$. By the monotonicity of T,

$$\langle u_n - x, w_n - y\rangle \geq 0$$

Since

$$\langle u_n, w_n\rangle = \langle u_n - x, w_n - y\rangle + \langle x, w_n\rangle + \langle u_n, y\rangle - \langle x, y\rangle,$$

where

$$\langle x, w_n\rangle + \langle u_n, y\rangle \to \langle x, w\rangle + \langle u, y\rangle - \langle x, y\rangle,$$

we have

$$\langle u, w\rangle \leq \limsup\langle u_n, w_n\rangle \geq \langle x, w\rangle + \langle u, y\rangle - \langle x, y\rangle,$$

i.e.,

$$\langle u - x, w - y\rangle \geq 0$$

Since T is maximal, $w \in T(u)$. Consequently,

$$\langle u_n - u, w_n - w \rangle \geq 0.$$

Thus,

$$\liminf \langle u_n, w_n \rangle \geq \liminf (\langle u_n, w \rangle \langle u, w_n \rangle + \langle u, w \rangle = \langle u, w \rangle,$$

It follows that $\lim \langle u_n, w_n \rangle \to \langle u, w \rangle$. $\square$

DEFINITION 5 (1) A mapping T from X into X^* is coercive if there exists a function $c : R+ \to R^+$ with $\lim_{r\infty} c(r) = \infty$ such that

$$\langle u, w \rangle \geq c(|u|)|u| \text{ for all } [u, w] \in G(T).$$

(1) An operator T from X into X^* is called smooth if it is bounded, coercive, maximal monotone, and has effective domain $D(T) = X$.
(2) Let T be a generalized pseudo-monotone mapping from X into X^*. Then T is said to be regular if $R(T + T_2) = X^*$ for each smooth operator T_2.
(3) T is said to be quasi-bounded if for each $M > 0$ there exists $K(M) > 0$ such that whenever $[u, w] \in G(T)$ of T and

$$\langle u, w \rangle \leq M\,|u|, \quad |u| \leq M,$$

then $|w| \leq K(M)$. T is said to be strongly quasi-bounded if for each $M > 0$ there exists $K(M) > 0$ such that whenever $[u, w] \in G(T)$ of T and

$$\langle u, w \rangle \leq M \quad |u| \leq M,$$

then $|w| \leq K(M)$.

Theorem 1 Let $T = T_1 + T_0$, where T_1 is maximal monotone with $0 \in D(T)$, while T_0 is a regular generalized pseudo-monotone mapping such that for a given constant k, $\langle u, w \rangle k\,|u|$ for all $[u, w] \in G(T)$. If either T_0 is quasi-bounded or T_1 is strongly quasi-bounded, then T is regular.

Theorem 2 (Browder). Let X be a reflexive Banach space and T be a generalized pseudo-monotone mapping from X into X^*. Suppose that $R(\epsilon F + T) = X^*$ for $\epsilon > 0$ and T is coercive. Then $R(T) = X^*$.

Lemma 1 Let T be a generalized pseudo-monotone mapping from the reflexive Banach space X into X^*, and let C be a bounded weakly closed subset of X. Then $T(C) = \{w \in X^* : w \in T(u) \text{for some } u \in C\}$ is closed in the strong topology of X^*.

Proof: Let Let $\{w_n\}$ be a sequence in $T(C)$ converging strongly to some $w \in X^*$. For each n, there exists $u_n \in C$ such that $w_n \in T(u_n)$. Since C is bounded and weakly closed, without loss of the generality we assume $u_n \to u \in C$, weakly in X. It follows that $\lim \langle u_n - u, w_n \rangle = 0$. The generalized pseudo-monotonicity of T implies that $w \in T(u)$, i.e., w lies in $T(C)$. $\square$

Proof of Theorem 2: Let J denote the normalized Let F denote the duality mapping of X into X^*. It is known that F is a smooth operator. Let w_0 be a given element of X^*. We wish to show that $w_0 \in R(T)$. For each $\epsilon > 0$, it follows from the regularity of T that there exists an element $u_\epsilon \in X$ such that

$$w_0 - \epsilon\, p_\epsilon \in T(u_\epsilon) \text{ for some } p_\epsilon \in F(u_\epsilon).$$

Let $w_\epsilon = w_0 - \epsilon\, p_\epsilon$. Since T is coercive and $\langle u_\epsilon, w_\epsilon\rangle = |u_\epsilon|^2$, we have for some k

$$\epsilon\, |u_\epsilon|^2 \leq k\, |u_\epsilon| + |u_\epsilon||w_0|.$$

Hence

$$\epsilon\, |u_\epsilon| = \epsilon\, |p_\epsilon| \leq |w_0|.$$

and

$$|w_\epsilon| \leq |w_0| + \epsilon\, |p_\epsilon| \leq k + 2|w_0|$$

Since T is coercive, there exits $M > 0$ such that $|u_\epsilon| \leq M$ and thus

$$|w_\epsilon - w_0| \leq \epsilon\, |p_\epsilon| = \epsilon\, |u_\epsilon| \to 0 \text{ as } \epsilon \to 0^+$$

It thus follows from Lemma 1 that $T(B(0, M))$ is closed and thus $w_0 \in T(B(0, M))$, i.e., $w_0 \in R(T)$. □

THEOREM 3 Let X be a reflexive Banach space and T a pseudo-monotone mapping from X into X^*. Suppose that T is coercive, $R(T) = X^*$.

PROPOSITION 10. Let F be a finite-dimensional Banach space, a mapping from F into F^* such that for each $u \in F$, T(u) is a nonempty bounded closed convex subset of F^*. Suppose that T is coercive and upper semicontinuous from F to F^*. Then $R(T) = F^*$.
Proof: Since for each $w_0 \in F^*$ define the mapping $T_{w_0} : F \to F^*$ by

$$T_{w_0}(u) = T(u) - w_0$$

satisfies the same hypotheses as the original mapping T. It thus suffices to prove that $0 \in R(T)$. Suppose that 0 does not lie in $R(T)$. Then for each $u \in F$, there exists $v(u) \in F$ such that

$$\inf_{w \in T(u)} \langle v(u), w\rangle > 0$$

Since T is coercive, there exists a positive constant R such that $c(R) > 0$, and hence for each $u \in F$ with $|u| = R$ and each $w \in T(u)$,

$$\langle u, w\rangle \geq \lambda > 0$$

where $\lambda = c(R)R$, For such points $u \in F$, we may take $v(u) = u$. For a given nonzero v_0 in F let

$$W_{v_0} = \{u \in F : \inf_{w \in T(u)} \langle v_0, w\rangle > 0\}.$$

The family $\{W_{v_0} : v_0 \in F, v_0 \neq 0\}$ forms a covering of the space F. By the upper semicontinuity of T, from F to F^* each W_{v_0} is open in F. Hence the family $\{W_{v_0} : v_0 \in F, v_0 \neq 0\}$ forms an open covering of F. We now choose a finite open covering $\{V_1, \cdots, V_m\}$ of the closed ball $B(0, R)$ in F of radius R about the origin, with the property that for each k there exists an element v_k in F such that $V_k \subset W_{v_k}$ and the additional condition that if V_k intersects the boundary sphere $S(0, R)$, then v_k is a point of $V_k \cap S(0, R)$ and the diameter of such V_k is less than $\frac{R}{2}$.. We next choose a partition of unity $\{\alpha_1, \cdots, \alpha_m\}$ subordinated to the covering $\{V_1, \cdots, V_m\}$ where each α_k is a continuous mapping of F into $[0, I]$ which vanishes outside the corresponding set V_k, and with the property that $\sum_k \alpha_k(x) = 1$ for each $x \in B(0, R)$. Using this partition of unity, we may define a continuous mapping f of $B(0, R)$ into F by setting

$$f(x) = \sum_k \alpha(x) v_k.$$

We note that for each k for which $alpha_k(x) > 0$ and for any $winT_0x$, we have $\langle v_k, w\rangle > 0$ since x must lie in V_k, which is a subset of W_k. Hence for each $x \in B(0, R)$ and any $w \in T_0x$

$$\langle f(x), w\rangle = \sum_k \alpha(x)\langle v_k, w\rangle > 0$$

Consequently, $f(x) \neq 0$ for any $x \in B(0, R)$. This implies that the degree of the mapping f on $B(0, R)$ with respect to 0 equals 0. For $x \in S(0, R)$, on the other hand, $f(x)$ is a convex linear combination of points $v_k \in S(0, R)$, each of which is at distance at most $\frac{R}{2}$ from the point x. We infer that

$$|f(x) - x| \leq \frac{R}{2}, \;\; x \in S(0, R).$$

By this inequality, f considered as a mapping from $B(0, R)$ into F is homotopic to the identity mapping. Hence the degree of f on $B(0, R)$ with respect to 0 is 1. This contradiction proves that 0 must lie in T_0u for some $u \in B(0, R)$. $\square$
Proof of Theorem 3 Let Λ be the family of all finite-dimensional subspaces F of X, ordered by inclusion. For $F \in \Lambda$, let j_F: $F \to X$ denote the inclusion mapping of F into X, and $j_F^* : X^* \to F^*$ the dual projection mapping of X^* onto F^*. The operator

$$T_F = j_F^* T j_F$$

then maps F into F^*. For each $u \in F$, $T_F u$ is a weakly compact convex subset of X^* and j_F^* is continuous from the weak topology on X^* to the (unique) topology on F^*. Hence $T_F u$ is a nonempty closed convex subset of F^*. Since T is upper semicontinuous from F to F^* with X^* given its weak topology, T_F is upper semicontinuous from F to F^*. By Proposition 10, to each $F \in \Lambda$ there exists an element $u_F \in F$ such that $j_F^* w_0 = w_F$ for some $w_F \in T_F u_F$. The coerciveness of T implies that $w_F \in T_F u_F$, i.e.

$$\langle u_F, j_F^* w_0\rangle = \langle u_F, w_F\rangle \geq c(|u_F|)|u_F|.$$

Consequently, the elements $\{|u_F|\}$ are uniformly bounded by M for all $F \in \Lambda$. For $F \in \Lambda$, let

$$V_F = \bigcup_{F' \in \Lambda,\, F' \subset F} \{u_{F'}\}.$$

Then the set V_F is contained in the closed ball $B(0, M)$ in X Since $B(0, M)$ is weakly compact, and since the family {weak closure(V_F)} has the finite intersection property, the intersection $\cap_{F \in \Lambda}${weak closure(V_F)} is not empty. Let u_0 be an element contained in this intersection.

The proof will be complete if we show that $0 \in T_{w_0} u_0$. Let $v \in X$ be arbitrarily given. We choose $F \in \Lambda$ such that it contains u_0, $v \in F$ Let $\{u_{F_k}\}$denote a sequence in V_F converging weakly to u_0. Since $j_{F_k}^* w_{F_k} = 0$, we have

$$\langle u_{F_k} - u_0, w_{F_k}\rangle = 0 \text{ for all } k.$$

Consequently, by the pseudo-monotonicity of T, to given $v \in X$ there exists $w(v) \in T(u_0)$ with

$$\lim \langle u_{F_k} - v, w_{F_k} \rangle \geq \langle u_0 - v, w \rangle. \tag{5.5}$$

Suppose now that $0 \notin T(u_0)$. Then 0 can be separated from the nonempty closed convex set $T(u_0)$, i.e., there exists an element $x = u - v \in X$ such that

$$\inf_{z \in T(u_0)} \langle u_0 - v, z \rangle > 0.$$

But this is a contradiction to (5.5). $\square$

Navier-Stokes equations

Example Consider the constrained minimization

$$\min \quad F(y) \text{ subject to } E(y) = 0.$$

where

$$E(y) = E_0 y + f(y).$$

The necessary optimality condition for $x = (y, p)$

$$A(y,p) \in \begin{pmatrix} \partial F - E'(y)^* p \\ E(y) \end{pmatrix}.$$

Let

$$A_0(y,p) \in \begin{pmatrix} \partial F - E_0^* p \\ E_0 y \end{pmatrix},$$

and

$$A_1(y,p) = \begin{pmatrix} -f'(y)p \\ f(y) \end{pmatrix}.$$

5.3 Dissipative Operators and Semigroup of Nonlinear Contractions

In this section we consider

$$\frac{du}{dt} \in Au(t), \quad u(0) = u_0 \in X$$

for the dissipative mapping A on a Banach space X.

Definition (Dissipative) A mapping A on a Banach space X is dissipative if

$$|x_1 - x_2 - \lambda\,(y_1 - y_2)| \geq |x_1 - x_2| \text{ for all } \lambda > 0 \text{ and } [x_1, y_1],\ [x_2, y_2] \in A,$$

or equivalently

$$\langle y_1 - y_2, x_1 - x_2 \rangle_- \leq 0 \text{ for all } [x_1, y_1],\ [x_2, y_2] \in A.$$

and if in addition $R(I - \lambda A) = X$, then A is m-dissipative.

In particular, it follows that if A is dissipative, then for $\lambda > 0$ the operator $(I - \lambda\, A)^{-1}$ is a single-valued and nonexpansive on $R(I - \lambda\, A)$, i.e.,

$$|(I - \lambda\, A)^{-1} x - (I - \lambda\, A)^{-1} y| \leq |x - y| \quad \text{for all } x,\ y \in R(I - \lambda\, A).$$

Define the resolvent and Yosida approximation A by

$$J_\lambda x = (I - \lambda\, A)^{-1} x \in dom\,(A), \quad x \in dom\,(J_\lambda) = R(I - \lambda\, A)$$
$$A_\lambda = \lambda^{-1}(J_\lambda x - x), \quad x \in dom\,(J_\lambda). \tag{5.6}$$

We summarize some fundamental properties of J_λ and A_λ in the following theorem.

Theorem 1.4 Let A be an $\omega-$ dissipative subset of $X \times X$, i.e.,

$$|x_1 - x_2 - \lambda\,(y_1 - y_2)| \geq (1 - \lambda\omega)\,|x_1 - x_2|$$

for all $0 < \lambda < \omega^{-1}$ and $[x_1, y_1],\ [x_2, y_2] \in A$ and define $\|Ax\|$ by

$$\|Ax\| = \inf\{|y| : y \in Ax\}.$$

Then for $0 < \lambda < \omega^{-1}$,

(i) $|J_\lambda x - J_\lambda y| \leq (1 - \lambda\omega)^{-1}\,|x - y|$ for $x,\ y \in dom\,(J_\lambda)$.
(ii) $A_\lambda x \in AJ_\lambda x$ for $x \in R(I - \lambda\, A)$.
(iii) For $x \in dom(J_\lambda) \cap dom\,(A)$ $|A_\lambda x| \leq (1 - \lambda\omega)^{-1}\,\|Ax\|$ and thus $|J_\lambda x - x| \leq \lambda(1 - \lambda\omega)^{-1}\,\|Ax\|$.
(iv) If $x \in dom\,(J_\lambda),\ \lambda,\ \mu > 0$, then

$$\frac{\mu}{\lambda}\,x + \frac{\lambda - \mu}{\lambda}\,J_\lambda x \in dom\,(J_\mu)$$

and

$$J_\lambda x = J_\mu\left(\frac{\mu}{\lambda}\,x + \frac{\lambda - \mu}{\lambda}\,J_\lambda x\right).$$

(v) If $x \in dom\,(J_\lambda) \cap dom\,(A)$ and $0 < \mu \leq \lambda < \omega^{-1}$, then $(1 - \lambda\omega)|A_\lambda x| \leq (1 - \mu\omega)\,|A_\mu y|$
(vi) A_λ is $\omega^{-1}(1 - \lambda\omega)^{-1}$-dissipative and for $x,\ y \in dom(J_\lambda)$ $|A_\lambda x - A_\lambda y| \leq \lambda^{-1}(1 + (1 - \lambda\omega)^{-1})\,|x - y|$.

Proof: $(i) - -(ii)$ If $x,\ y \in dom\,(J_\lambda)$ and we set $u = J_\lambda x$ and $v = J_\lambda y$, then there exist $\hat{u}$ and $\hat{v}$ such that $x = u - \lambda\,\hat{u}$ and $y = v - \lambda\,\hat{v}$. Thus, from (1.8)

$$|J_\lambda x - J_\lambda y| = |u - v| \leq (1 - \lambda\omega)^{-1}|u - v - \lambda\,(\hat{u} - \hat{v})| = (1 - \lambda\omega)^{-1}\,|x - y|.$$

Next, by the definition $A_\lambda x = \lambda^{-1}(u - x) = \hat{u} \in Au = AJ_\lambda x$.
(iii) Let $x \in dom\,(J_\lambda) \cap dom\,(A)$ and $\hat{x} \in Ax$ be arbitrary. Then we have $J_\lambda(x - \lambda\,\hat{x}) = x$ since $x - \lambda\,\hat{x} \in (I - \lambda\,A)x$. Thus,

$$|A_\lambda x| = \lambda^{-1}|J_\lambda x - x| = \lambda^{-1}|J_\lambda x - J_\lambda(x - \lambda\,\hat{x})| \leq (1 - \lambda\omega)^{-1}\lambda^{-1}\,|x - (x - \lambda\,\hat{x})| = (1 - \lambda\omega)^{-1}\,|\hat{x}|.$$

which implies (iii).
(iv) If $x \in dom\,(J_\lambda) = R(I - \lambda\,A)$ then we have $x = u - \lambda\,\hat{u}$ for $[u, \hat{u}] \in A$ and thus $u = J_\lambda x$. For $\mu > 0$

$$\frac{\mu}{\lambda}\,x + \frac{\lambda - \mu}{\lambda}\,J_\lambda x = \frac{\mu}{\lambda}\,(u - \lambda\,\hat{u}) + \frac{\lambda - \mu}{\lambda}\,u = u - \mu\,\hat{u} \in R(I - \mu\,A) = dom\,(J_\mu).$$

and

$$J_\mu\left(\frac{\mu}{\lambda}\,x + \frac{\lambda - \mu}{\lambda}\,J_\lambda x\right) = J_\mu(u - \mu\,\hat{u}) = u = J_\lambda x.$$

(v) From (i) and (iv) we have

$$\lambda\,|A_\lambda x| = |J_\lambda x - x| \le |J_\lambda x - J_\lambda y| + |J_\mu x - x|$$

$$\le \left| J_\mu \left(\frac{\mu}{\lambda}\, x + \frac{\lambda-\mu}{\lambda}\, J_\lambda x \right) - J_\mu x \right| + |J_\mu x - x|$$

$$\le (1-\mu\omega)^{-1} \left| \frac{\mu}{\lambda}\, x + \frac{\lambda-\mu}{\lambda}\, J_\lambda x - x \right| + |J_\mu x - x|$$

$$= (1-\mu\omega)^{-1}(\lambda-\mu)\,|A_\lambda x| + \mu\,|A_\mu x|,$$

which implies (v) by rearranging.
(vi) It follows from (i) that for $\rho > 0$

$$|x - y - \rho\,(A_\lambda x - A_\lambda y)| = |(1+\frac{\rho}{\lambda})\,(x-y) - \frac{\rho}{\lambda}\,(J_\lambda x - J_\lambda y)|$$

$$\ge (1+\frac{\rho}{\lambda})\,|x-y| - \frac{\rho}{\lambda}\,|J_\lambda x - J_\lambda y|$$

$$\ge ((1+\frac{\rho}{\lambda}) - \frac{\rho}{\lambda}(1-\lambda\omega)^{-1})\,|x-y| = (1-\rho\,\omega(1-\lambda\omega)^{-1})\,|x-y|.$$

The last assertion follows from the definition of A_λ and (i). $\square$

Theorem 1.5
(1) A dissipative set $A \subset X \times X$ is m-dissipative, if and only if

$$R(I - \lambda_0\,A) = X \quad \text{for some } \ \lambda_0 > 0.$$

(2) An m-dissipative mapping is maximal dissipative, i.e., all dissipative set containing A in $X \times X$ coincide with A.
(3) If $X = X^* = H$ is a Hilbert space, then the notions of the maximal dissipative set and m-dissipative set are equivalent.

Proof: (1) Suppose $R(I - \lambda_0\,A) = X$. Then it follows from Theorem 1.4 (i) that J_{λ_0} is contraction on X. We note that

$$I - \lambda\,A = \frac{\lambda_0}{\lambda} \left(I - (1 - \frac{\lambda_0}{\lambda})\,J_{\lambda_0} \right) (I - \lambda_0\,A) \tag{5.7}$$

for $0 < \lambda < \omega^{-1}$. For given $x \in X$ define the operator $T : X \to X$ by

$$Ty = x + (1 - \frac{\lambda_0}{\lambda})\,J_{\lambda_0} y, \quad y \in X$$

Then

$$|Ty - Tz| \le |1 - \frac{\lambda_0}{\lambda}|\,|y-z|$$

where $|1 - \frac{\lambda}{\lambda_0}| < 1$ if $2\lambda > \lambda_0$. By Banach fixed-point theorem the operator T has a unique fixed point $z \in X$, i.e., $x = (I - (1 - \frac{\lambda_0}{\lambda}\,J_{\lambda_0})z$. Thus,

$$x \in (I - (1 - \frac{\lambda_0}{\lambda})\,J_{\lambda_0})(I - \lambda_0\,A)dom\,(A).$$

and it thus follows from (5.7) that $R(I-\lambda\, A)=X$ if $\lambda>\frac{\lambda_0}{2}$. Hence, (1) follows from applying the above argument repeatedly.
(2) Assume A is m-dissipative. Suppose $\tilde{A}$ is a dissipative set containing A. We need to show that $\tilde{A}\subset A$. Let $[x,\hat{x}]\in\tilde{A}$ Since $x-\lambda\,\hat{x}\in X=R(I-\lambda\,A)$, for $\lambda>0$, there exists a $[y,\hat{y}]\in A$ such that $x-\lambda\,\hat{x}=y-\lambda\,\hat{y}$. Since $A\subset\tilde{A}$ it follows that $[y,\hat{y}]\in A$ and thus

$$|x-y|\leq|x-y-\lambda\,(\hat{x}-\hat{y})|=0.$$

Hence, $[x,\hat{x}]=[y,\hat{y}]\in A$.
(3) It suffices to show that if A is maximal dissipative, then A is m-dissipative. We use the following extension lemma, Lemma 1.6. Let y be any element of H. By Lemma 1.6, taking $C=H$, we have that there exists $x\in H$ such that

$$(\xi-x,\eta-x+y)\leq 0\quad\text{for all }\ [\xi,\eta]\in A.$$

and thus

$$(\xi-x,\eta-(x-y))\leq 0\quad\text{for all }\ [\xi,\eta]\in A.$$

Since A is maximal dissipative, this implies that $[x,x-y]\in A$, that is $x-y\in Ax$, and therefore $y\in R(I-H)$. $\square$

<u>Lemma 1.6</u> Let A be dissipative and C be a closed, convex, non-empty subset of H such that $dom\,(A)\in C$. Then for every $y\in H$ there exists $x\in C$ such that

$$(\xi-x,\eta-x+y)\leq 0\quad\text{for all }\ [\xi,\eta]\in A.$$

Proof: Without loss of generality we can assume that $y=0$, for otherwise we define $A_y=\{[\xi,\eta+y]:[\xi,\eta]\in A\}$ with $dom\,(A_y)=dom\,(A)$. Since A is dissipative if and only if A_y is dissipative , we can prove the lemma for A_y. For $[\xi,\eta]\in A$, define the set

$$C([\xi,\eta])=\{x\in C:(\xi-x,\eta-x)\leq 0\}.$$

Thus, the lemma is proved if we can show that $\bigcap_{[\xi,\eta]\in A}\,C([\xi,\eta])$ is non-empty.

5.3.1 Properties of m-dissipative operators

In this section we discuss some properties of m-dissipative sets.

<u>Lemma 1.7</u> Let X^* be a strictly convex Banach space. If A is maximal dissipative, then Ax is a closed convex set of X for each $x\in dom\,(A)$.

Proof: It follows from Lemma that the duality mapping F is single-valued. First, we show that Ax is convex. Let $\hat{x}_1,\ \hat{x}_2\in Ax$ and set $\hat{x}=\alpha\hat{x}_1+(1-\alpha)\,\hat{x}_2$ for $0\leq\alpha\leq 1$. Then, Since A is dissipative, for all $[y,\hat{y}]\in A$

$$Re\,\langle\hat{x}-\hat{y},F(x-y)\rangle=\alpha\,Re\,\langle\hat{x}_1-\hat{y},F(x-y)\rangle+(1-\alpha)\,Re\,\langle\hat{x}_2-\hat{y},F(x-y)\rangle\leq 0.$$

Thus, if we define a subset $\tilde{A}$ by

$$\tilde{A}z=\begin{cases}Az & \text{if }\ z\in dom\,(A)\setminus\{x\}\\[2ex] Ax\cup\{\hat{x}\} & \text{if }\ z=x,\end{cases}$$

then $\tilde{A}$ is a dissipative extension of A and $dom\,(\tilde{A}) = dom\,(A)$. Since A is maximal dissipative, it follows that $\tilde{A}x = Ax$ and thus $\hat{x} \in Ax$ as desired.

Next, we show that Ax is closed. Let $\hat{x}_n \in Ax$ and $\lim_{n\to\infty} \hat{x}_n = \hat{x}$. Since A is dissipative, $Re\,\langle \hat{x}_n - \hat{y}, x - y\rangle \leq 0$ for all $[y, \hat{y}] \in A$. Letting $n \to \infty$, we obtain $Re\,\langle \hat{x} - \hat{y}, x - y\rangle \leq 0$. Hence, as shown above $\hat{x} \in Ax$ as desired. $\square$

Definition 1.4 A subset A of $X \times X$ is said to be demiclosed if $x_n \to x$ and $y_n \rightharpoonup y$ and $[x_n, y_n] \in A$ imply that $[x, y] \in A$. A subset A is closed if $[x_n, y_n]$, $x_n \to x$ and $y_n \to y$ imply that $[x, y] \in A$.

Theorem 1.8 Let A be m-dissipative. Then the followings hold.
(i) A is closed.
(ii) If $\{x_\lambda\} \subset X$ such that $x_\lambda \to x$ and $A_\lambda x_\lambda \to y$ as $\lambda \to 0^+$, then $[x, y] \in A$.

Proof: (i) Let $[x_n, \hat{x}_n] \in A$ and $(x_n, \hat{x}_n) \to (x, \hat{x})$ in $X \times X$. Since A is dissipative $Re\,\langle \hat{x}_n - \hat{y}, x_n - y\rangle_i \leq 0$ for all $[y, \hat{y}] \in A$. Since $\langle \cdot, \cdot\rangle_i$ is lower semicontinuous, letting $n \to \infty$, we obtain $Re\,\langle \hat{x} - \hat{y}, x - y\rangle_i \leq$ for all $[y, \hat{y}] \in A$. Then $A_1 = [x, \hat{x}] \cup A$ is a dissipative extension of A. Since A is maximal dissipative, $A_1 = A$ and thus $[x, \hat{x}] \in A$. Hence, A is closed.
(ii) Since $\{A_\lambda x\}$ is a bounded set in X, by the definition of A_λ, $\lim\, |J_\lambda x_\lambda - x_\lambda| \to 0$ and thus $J_\lambda x_\lambda \to x$ as $\lambda \to 0^+$. But, since $A_\lambda x_\lambda \in AJ_\lambda x_\lambda$, it follows from ($i$) that $[x, y] \in A$. $\square$

Theorem 1.9 Let A be m-dissipative and let X^* be uniformly convex. Then the followings hold.
(i) A is demiclosed.
(ii) If $\{x_\lambda\} \subset X$ such that $x_\lambda \to x$ and $\{|A_\lambda x|\}$ is bounded as $\lambda \to 0^+$, then $x \in dom\,(A)$. Moreover, if for some subsequence $A_{\lambda_n} x_n \rightharpoonup y$, then $y \in Ax$.
(iii) $lim_{\lambda\to 0^+} |A_\lambda x| = \|Ax\|$.

Proof: (i) Let $[x_n, \hat{x}_n] \in A$ be such that $\lim\, x_n = x$ and $w - \lim \hat{x}_n = \hat{x}$ as $n \to \infty$. Since X^* is uniformly convex, from Lemma the duality mapping is single-valued and uniformly continuous on the bounded subsets of X. Since A is dissipative $Re\,\langle \hat{x}_n - \hat{y}, F(x_n - y)\rangle \leq 0$ for all $[y, \hat{y}] \in A$. Thus, letting $n \to \infty$, we obtain $Re\,\langle \hat{x} - \hat{y}, F(x - y)\rangle \leq 0$ for all $[y, \hat{y}] \in A$. Thus, $[x, \hat{x}] \in A$, by the maximality of A.
Definition 1.5 The minimal section A^0 of A is defined by

$$A^0x = \{y \in Ax : |y| = \|Ax\|\} \quad \text{with } \; dom\,(A^0) = \{x \in dom\,(A) : A^0x \text{ is non-empty}\}.$$

Lemma 1.10 Let X^* be a strictly convex Banach space and let A be maximal dissipative. Then, the followings hold.
(i) If X is strictly convex, then A^0 is single-valued.
(ii) If X reflexible, then $dom\,(A^0) = dom\,(A)$.
(iii) If X strictly convex and reflexible, then A^0 is single-valued and $dom\,(A^0) = dom\,(A)$.

Theorem 1.11 Let X^* is a uniformly convex Banach space and let A be m-dissipative. Then the followings hold.
(i) $lim_{\lambda\to 0^+} F(A_\lambda x) = F(A^0 x)$ for each $x \in dom\,(A)$.
Moreover, if X is also uniformly convex, then
(ii) $\lim_{\lambda\to 0^+} A_\lambda x = A^0 x$ for each $x \in dom\,(A)$.

Proof: (1) Let $x \in dom\,(A)$. By (ii) of Theorem 1.2

$$|A_\lambda x| \le \|Ax\|$$

Since $\{A_\lambda x\}$ is a bounded sequence in a reflexive Banach space (i.e., since X^* is uniformly convex, X^* is reflexive and so is X), there exists a weak convergent subsequence $\{A_{\lambda_n} x\}$. Now we set $y = w - \lim_{n\to\infty} A_{\lambda_n} x$. Since from Theorem 1.2 $A_{\lambda_n} x \in AJ_{\lambda_n} x$ and $\lim_{n\to\infty} J_{\lambda_n} x = x$ and from Theorem 1.10 A is demiclosed, it follows that $[x, y] \in A$. Since by the lower-semicontinuity of norm this implies

$$\|Ax\| \le |y| \liminf_{n\to\infty} |A_{\lambda_n} x| \le \limsup_{n\to\infty} |A_{\lambda_n} x| \le \|Ax\|,$$

we have $|y| = \|Ax\| = \lim_{n\to\infty} |A_{\lambda_n} x|$ and thus $y \in A^0 x$. Next, since $|F(A_{\lambda_n} x)| = |A_{\lambda_n} x| \le \|Ax\|$, $F(A_{\lambda_n} x)$ is a bounded sequence in the reflexive Banach space X^* and has a weakly convergent subsequence $F(A_{\lambda_k} x)$ of $F(A_{\lambda_n} x)$. If we set $y^* = w - \lim_{k\to\infty} F(A_{\lambda_k} x)$, then it follows from the dissipativity of A that

$$Re\,\langle A_{\lambda_k} x - y, F(A_{\lambda_k} x)\rangle = \lambda_n^{-1} Re\,\langle A_{\lambda_k} x - y, F(J_{\lambda_k} x x)\rangle \le 0,$$

or equivalently $|A_{\lambda_k} x|^2 \le Re\,\langle y, F(A_{\lambda_n} x)\rangle$. Letting $k \to \infty$, we obtain $|y|^2 \le Re\,\langle y, y^*\rangle$. Combining this with

$$|y^*| \le \lim_{k\to\infty} |F(A_{\lambda_k} x)| = \lim_{k\to\infty} |A_{\lambda_k} x| = |y|,$$

we have

$$|y^*| \le Re\,\langle y, y^*\rangle \le |\langle y, y^*\rangle| \le |y||y^*| \le |y|^2.$$

Hence,

$$\langle y, y^*\rangle = |y|^2 = |y^*|^2$$

and we have $y^* = F(y)$. Also, $\lim_{k\to\infty} |F(A_{\lambda_k} x)| = |y| = |F(y)|$. It thus follows from the uniform convexity of X^* that

$$\lim_{k\to\infty} F(A_{\lambda_k} x) = F(y).$$

Since Ax is a closed convex set of X from Theorem , we can show that $x \to F(A^0 x)$ is single-valued. In fact, if C is a closed convex subset of X, the y is an element of minimal norm in C, if and only if

$$|y| \le |(1-\alpha)y + \alpha\, z| \quad \text{for all } z \in C \text{ and } 0 \le \alpha \le 1.$$

Hence,

$$\langle z - y, y\rangle_+ \ge 0.$$

and from Theorem 1.10

$$0 \le Re\,\langle z - y, f\rangle = Re\,\langle z, f\rangle - |y|^2 \tag{5.8}$$

for all $z \in C$ and $f \in F(y)$. Now, let $y_1,\ y_2$ be arbitrary in $A^0 x$. Then, from (5.8)

$$|y_1|^2 \le Re\,\langle y_2, F(y_1)\rangle \le |y_1||y_2|$$

which implies that $\langle y_2, F(y_1)\rangle = |y_2|^2$ and $|F(y_1)| = |y_2|$. Therefore, $F(y_1) = F(y_2)$ as desired. Thus, we have shown that for every sequence $\{\lambda\}$ of positive numbers that converge to zero, the sequence $\{F(A_\lambda x)\}$ has a subsequence that converges to the same limit $F(A^0x)$. Therefore, $\lim_{\lambda\to 0} F(A_\lambda x) \to F(A^0x)$.

Furthermore, we assume that X is uniformly convex. We have shown above that for $x \in dom\,(A)$ the sequence $\{A_\lambda\}$ contains a weak convergent subsequence $\{A_{\lambda_n}x\}$ and if $y = w - \lim_{n\to\infty} A_{\lambda_n}x$ then $[x, y] \in A^0$ and $|y| = \lim_{n\to\infty} |A_{\lambda_n}x|$. But since X is uniformly convex, it follows from Theorem 1.10 that A^0 is single-valued and thus $y = A^0x$. Hence, $w - \lim_{n\to\infty} A_{\lambda_n}x = A^0x$ and $\lim_{n\to\infty} |A_{\lambda_n}x| = |A^0x|$. Since X is uniformly convex, this implies that $\lim_{n\to\infty} A_{\lambda_n}x = A^0x$. $\square$

Theorem 1.12 Let X is a uniformly convex Banach space and let A be m-dissipative. Then $\overline{dom\,(A)}$ is a convex subset of X.

Proof: It follows from Theorem 1.4 that

$$|J_\lambda x - x| \le \lambda\,\|Ax\| \quad \text{for } \ x \in dom\,(A)$$

Hence $|J_\lambda x - x| \to 0$ as $\lambda \to 0^+$. Since $J_\lambda x \in dom\,(A)$ for $X \in X$, it follows that

$$\overline{dom\,(A)} = \{x \in X : |J_\lambda x - x| \to 0 \text{ as } \lambda \to 0^+\}.$$

Let $x_1,\ x_2 \in \overline{dom\,(A)}$ and $0 \le \alpha \le 1$ and set

$$x = \alpha\, x_1 + (1-\alpha)\, x_2.$$

Then, we have

$$\begin{aligned} |J_\lambda x - x_1| &\le |x - x_1| + |J_\lambda x_1 - x_1| \\ |J_\lambda x - x_2| &\le |x - x_2| + |J_\lambda x_2 - x_2| \end{aligned} \tag{5.9}$$

where $x - x_1 = (1-\alpha)\,(x_2 - x_1)$ and $x - x_2 = \alpha\,(x_1 - x_2)$. Since $\{J_\lambda x\}$ is a bounded set and a uniformly convex Banach space is reflexive, it follows that there exists a subsequence $\{J_{\lambda_n}x\}$ that converges weakly to z. Since the norm is weakly lower semi-continuous, letting $n \to \infty$ in (5.9) with $\lambda = \lambda_n$, we obtain

$$\begin{aligned} &|z - x_1| \le (1-\alpha)\,|x_1 - x_2| \\ &|z - x_2| \le \alpha\,|x_1 - x_2| \end{aligned}$$

Thus,

$$|x_1 - x_2| = |(x_1 - z) + (z - x_2)| \le |x_1 - z| + |z - x_2| \le |x_1 - x_2|$$

and therefore $|x_1 - z| = (1-\alpha)\,|x_1 - x_2|$, $|z - x_2| = \alpha\,|x_1 - x_2|$ and $|(x_1 - z) + (z - x_2)| = |x_1 - x_2|$. But, since X is uniformly convex we have $z = x$ and $w - \lim J_{\lambda_n}x = x$ as $n \to \infty$. Since we also have

$$|x - x_1| \le \liminf_{n\to\infty} |J_{\lambda_n}x - J_{\lambda_n}x_1| \le |x - x_1|$$

$|J_{\lambda_n}x - J_{\lambda_n}x_1| \to |x - x_1|$ and $w - \lim J_{\lambda_n}x - J_{\lambda_n}x_1 = x - x_1$ as $n \to \infty$. Since X is uniformly convex, this implies that $\lim_{\lambda\to 0^+} J_\lambda x = x$ and $x \in \overline{dom\,(A)}$. $\square$

5.3.2 Generation of Nonlinear Semigroups

Definition 2.1 Let X_0 be a subset of X. A semigroup $S(t),\ t \geq$ of nonlinear contractions on X_0 is a function with domain $[0,\infty) \times X_0$ and range in X_0 satisfying the following conditions:

$$S(t+s)x = S(t)S(s)x \quad \text{and} \quad S(0)x = x \quad \text{for } x \in X_0,\ t,s \geq 0$$

$$t \to S(t)x \in X \ \text{ is continuous}$$

$$|S(t)x - S(t)y| \leq |x-y| \quad \text{for } t \geq 0,\ x,y \in X_0$$

In this section, we consider the generation of nonlinear semigroup by Crandall-Liggett on a Banach space X. Let A be a ω-dissipative operator and $J_\lambda = (I - \lambda\, A)^{-1}$ is the resolvent. The following estimate plays an essential role in the Crandall-Liggett generation theory.

Lemma 2.1 Assume a sequence $\{a_{n,m}\}$ of positive numbers satisfies

$$a_{n,m} \leq \alpha\, a_{n-1,m-1} + (1-\alpha)\, a_{n-1,m} \tag{5.10}$$

and $a_{0,m} \leq m\,\lambda$ and $a_{n,0} \leq n\,\mu$ for $\lambda \geq \mu > 0$ and $\alpha = \frac{\mu}{\lambda}$. Then we have the estimate

$$a_{n,m} \leq [(m\lambda - n\mu)^2 + m\lambda^2]^{\frac{1}{2}} + [(m\lambda - n\mu)^2 + n\lambda\mu]^{\frac{1}{2}}. \tag{5.11}$$

Proof: From the assumption, (5.11) holds for either $m = 0$ or $n = 0$. We will use the induction in $n,\ m$, that is if (5.11) holds for $(n+1,m)$ when (5.11) is true for (n,m) and $(n,m-1)$, then (5.11) holds for all (n,m). Let $\beta = 1-\alpha$. We assume that (5.11) holds for (n,m) and $(n,m-1)$. Then, by (5.10) and Cauchy-Schwarz inequality $\alpha x + \beta y \leq (\alpha+\beta)^{\frac{1}{2}}(\alpha x^2 + \beta y^2)^{\frac{1}{2}}$

$$\begin{aligned}
a_{n+1,m} &\leq \alpha\, a_{n,m-1} + \beta\, a_{n,m} \\
&\leq \alpha\,([((m-1)\lambda - n\mu)^2 + (m-1)\lambda^2]^{\frac{1}{2}} + [((m-1)\lambda - n\mu)^2 + n\lambda\mu]^{\frac{1}{2}}) \\
&\quad +\beta\,([(m\lambda - n\mu)^2 + m\lambda^2]^{\frac{1}{2}} + [(m\lambda - n\mu)^2 + n\lambda\mu]^{\frac{1}{2}}) \\
&= (\alpha+\beta)^{\frac{1}{2}}(\alpha\,[((m-1)\lambda - n\mu)^2 + (m-1)\lambda^2] + \beta\,[(m\lambda - n\mu)^2 + m\lambda^2])^{\frac{1}{2}} \\
&\quad +(\alpha+\beta)^{\frac{1}{2}}(\alpha\,[((m-1)\lambda - n\mu)^2 + n\lambda\mu] + \beta\,[(m\lambda - n\mu)^2 + n\lambda\mu])^{\frac{1}{2}} \\
&\leq [(m\lambda - (n+1)\mu)^2 + m\lambda^2]^{\frac{1}{2}} + [(\lambda - (n+1)\mu)^2 + (n+1)\lambda\mu]^{\frac{1}{2}}.
\end{aligned}$$

Here, we used $\alpha + \beta = 1$, $\alpha\lambda = \mu$ and

$$\begin{aligned}
&\alpha\,[((m-1)\lambda - n\mu)^2 + (m-1)\lambda^2] + \beta\,[(m\lambda - n\mu)^2 + m\lambda^2]) \\
&\qquad \leq (m\lambda - n\mu)^2 + m\lambda^2 - \alpha\lambda(m\lambda - n\mu) = (m\lambda - (n+1)\mu)^2 + m\lambda^2 - \mu^2
\end{aligned}$$

$$\begin{aligned}
&\alpha\,[((m-1)\lambda - n\mu)^2 + n\lambda\mu] + \beta\,[(m\lambda - n\mu)^2 + n\lambda\mu] \\
&\qquad \leq (m\lambda - n\mu)^2 + (n+1)\lambda\mu - 2\alpha\lambda(m\lambda - n\mu) \leq (m\lambda - (n+1)\mu)^2 + (n+1)\lambda\mu - \mu^2.\square
\end{aligned}$$

Theorem 2.2 Assume A be a dissipative subset of $X \times X$ and satisfies the range condition

$$\overline{dom\,(A)} \subset R(I - \lambda\, A) \quad \text{for all sufficiently small } \lambda > 0. \tag{5.12}$$

Then, there exists a semigroup of type ω on $S(t)$ on $\overline{dom\,(A)}$ that satisfies for $x \in \overline{dom\,(A)}$

$$S(t)x = \lim_{\lambda \to 0^+} (I - \lambda\, A)^{-[\frac{t}{\lambda}]}x, \quad t \geq 0 \tag{5.13}$$

and

$$|S(t)x - S(t)x| \leq |t - s|\, \|Ax\| \quad \text{for } x \in dom\,(A), \text{ and } t,\ s \geq 0.$$

Proof: First, note that from (5.14) $\overline{dom\,(A)} \subset dom\,(J_\lambda)$. Let $x \in dom\,(A)$ and set $a_{n,m} = |J_\mu^n x - J_\lambda^m x|$ for $n,\ m \geq 0$. Then, from Theorem 1.4

$$a_{0,m} = |x - J_\lambda^m x| \leq |x - J_\lambda x| + |J_\lambda x - J_\lambda^2 x| + \cdots + |J_\lambda^{m-1}x - J_\lambda^m x|$$

$$\leq m\, |x - J_\lambda x| \leq m\lambda\, \|Ax\|.$$

Similarly, $a_{n,0} = |J_\mu^n x - x| \leq n\mu\, \|Ax\|$. Moreover,

$$a_{n,m} = |J_\mu^n x - J_\lambda^m x| \leq |J_\mu^n x - J_\mu \left(\frac{\mu}{\lambda}\, J_\lambda^{m-1}x + \frac{\lambda - \mu}{\lambda}\, J_\lambda^m x\right)|$$

$$\leq \frac{\mu}{\lambda}\, |J_\mu^{n-1}x - J_\lambda^{m-1}x| + \frac{\lambda - \mu}{\lambda}\, |J_\mu^{n-1}x - J_\lambda^m x|$$

$$= \alpha\, a_{n-1,m-1} + (1 - \alpha)\, a_{n-1,m}.$$

It thus follows from Lemma 2.1 that

$$|J_\mu^{[\frac{t}{\mu}]}x - J_\lambda^{[\frac{t}{\lambda}]}x| \leq 2(\lambda^2 + \lambda t)^{\frac{1}{2}}\, \|Ax\|. \tag{5.14}$$

Thus, $J_\lambda^{[\frac{t}{\lambda}]}x$ converges to $S(t)x$ uniformly on any bounded intervals, as $\lambda \to 0^+$. Since $J_\lambda^{[\frac{t}{\lambda}]}$ is non-expansive, so is $S(t)$. Hence (5.13) holds. Next, we show that $S(t)$ satisfies the semigroup property $S(t+s)x = S(t)S(s)x$ for $x \in \overline{dom\,(A)}$ and $t,\ s \geq 0$. Letting $\mu \to 0^+$ in (**??**), we obtain

$$|T(t)x - J_\lambda^{[\frac{t}{\lambda}]}x| \leq 2(\lambda^2 + \lambda t)^{\frac{1}{2}}. \tag{5.15}$$

for $x \in dom\,(A)$. If we let $x = J_\lambda^{[\frac{s}{\lambda}]}z$, then $x \in dom\,(A)$ and

$$|S(t)J_\lambda^{[\frac{s}{\lambda}]}z - J_\lambda^{[\frac{t}{\lambda}]}J_\lambda^{[\frac{s}{\lambda}]}z| \leq 2(\lambda^2 + \lambda t)^{\frac{1}{2}}\, \|Az\| \tag{5.16}$$

where we used that $\|AJ_\lambda^{[\frac{s}{\lambda}]}z\| \leq \|Az\|$ for $z \in dom\,(A)$. Since $[\frac{t+s}{\lambda}] - ([\frac{t}{\lambda}] + [\frac{s}{\lambda}])$ equals 0 or 1, we have

$$|J_\lambda^{[\frac{t+s}{\lambda}]}z - J_\lambda^{[\frac{t}{\lambda}]}J_\lambda^{[\frac{s}{\lambda}]}z| \leq |J_\lambda z - z| \leq \lambda\, \|Az\|. \tag{5.17}$$

It thus follows from (5.15)–(5.17) that

$$|S(t+s)z - S(t)S(s)z| \leq |S(t+s)x - J_\lambda^{[\frac{t+s}{\lambda}]}z| + |J_\lambda^{[\frac{t+s}{\lambda}]}z - J_\lambda^{[\frac{t}{\lambda}]}J_\lambda^{[\frac{s}{\lambda}]}z|$$

$$+|J_\lambda^{[\frac{t}{\lambda}]}J_\lambda^{[\frac{s}{\lambda}]}z - S(t)J_\lambda^{[\frac{s}{\lambda}]}z| + |S(t)J_\lambda^{[\frac{s}{\lambda}]}z - S(t)S(s)z| \to 0$$

as $\lambda \to 0^+$. Hence, $S(t+s)z = S(t)S(s)z$.

Finally, since

$$|J_\lambda^{[\frac{t}{\lambda}]}x - x| \leq [\frac{t}{\lambda}]\,|J_\lambda x - x| \leq t\,\|Ax\|$$

we have $|S(t)x - x| \leq t\,\|Ax\|$ for $x \in dom\,(A)$. From this we obtain $|S(t)x - x| \to 0$ as $t \to 0^+$ for $x \in \overline{dom\,(A)}$ and also

$$|S(t)x - S(s)x| \leq |S(t-s)x - x| \leq (t-s)\,\|Ax\|$$

for $x \in dom\,(A)$ and $t \geq s \geq 0$. $\square$

6 Cauchy Problem

Definition 3.1 Let $x_0 \in X$ and $\omega \in R$. Consider the Cauchy problem

$$\frac{d}{dt}u(t) \in Au(t), \quad u(0) = x_0. \tag{6.1}$$

(1) A continuous function $u(t) : [0,T] \to X$ is called a strong solution of (6.1) if $u(t)$ is Lipschitz continuous with $u(0) = x_0$, strongly differentiable a.e. $t \in [0,T]$, and (6.1) holds a.e. $t \in [0,T]$.
(2) A continuous function $u(t) : [0,T] \to X$ is called an integral solution of type ω of (6.1) if $u(t)$ satisfies

$$|u(t) - x| - |u(\hat{t}) - x| \leq \int_{\hat{t}}^t (\omega\,|u(s) - x| + \langle y, u(s) - x\rangle_+)\,ds \tag{6.2}$$

for all $[x,y] \in A$ and $t \geq \hat{t} \in [0,T]$.

Theorem 3.1 Let A be a dissipative subset of $X \times X$. Then, the strong solution to (6.1) is unique. Moreover if the range condition (5.14) holds, then then the strong solution $u(t) : [0,\infty) \to X$ to (6.1) is given by

$$u(t) = \lim_{\lambda \to 0^+} (I - \lambda\,A)^{-[\frac{t}{\lambda}]}x \quad \text{for } x \in \overline{dom\,(A)} \text{ and } t \geq 0.$$

Proof: Let $u_i(t),\ i = 1,2$ be the strong solutions to (6.1). Then, $t \to |u_1(t) - u_2(t)|$ is Lipschitz continuous and thus a.e. differentiable $t \geq 0$. Thus

$$\frac{d}{dt}|u_1(t) - u_2(t)|^2 = 2\,\langle u_1'(t) - u_2'(t), u_1(t) - u_2(t)\rangle_i$$

a.e. $t \geq 0$. Since $u_i'(t) \in Au_i(t),\ i = 1,2$ from the dissipativeness of A, we have $\frac{d}{dt}|u_1(t) - u_2(t)|^2 \leq 0$ and therefore

$$|u_1(t) - u_2(t)|^2 \leq \int_0^t \frac{d}{dt}|u_1(t) - u_2(t)|^2\,dt \leq 0,$$

which implies $u_1 = u_2$.

For $0 < 2\lambda < s$ let $u_\lambda(t) = (I - \lambda\,A)^{-[\frac{t}{\lambda}]}x$ and define $g_\lambda(t) = \lambda^{-1}(u(t) - u(t-\lambda)) - u'(t)$ a.e. $t \geq \lambda$. Since $\lim_{\lambda \to 0^+} |g_\lambda| = 0$ a.e. $t > 0$ and $|g_\lambda(t)| \leq 2M$ for a.e. $t \in [\lambda, s]$, where M is a Lipschitz constant of $u(t)$ on $[0,s]$, it follows that $\lim_{\lambda\to 0^+}\int_\lambda^s |g_\lambda(t)|\,dt = 0$ by Lebesgue dominated convergence theorem. Next, since

$$u(t-\lambda)+\lambda\, g_\lambda(t) = u(t)-\lambda\, u'(t) \in (I-\lambda\, A)u(t),$$

we have $u(t)=(I-\lambda\, A)^{-1}(u(t-\lambda)+\lambda\, g_\lambda(t))$. Hence,

$$|u_\lambda(t)-u(t)| \le |(I-\lambda\, A)^{-[\frac{t-\lambda}{\lambda}]}x - u(t-\lambda)-\lambda\, g_\lambda(t)|$$

$$\le |u_\lambda(t-\lambda)-u(t-\lambda)|+\lambda\,|g_\lambda(t)|$$

a.e. $t\in[\lambda,s]$. Integrating this on $[\lambda,s]$, we obtain

$$\lambda^{-1}\int_{s-\lambda}^s |u_\lambda(t)-u(t)|\,dt \le \lambda^{-1}\int_0^\lambda |u_\lambda(t)-u(t)|\,dt + \int_\lambda^s |g_\lambda(t)|\,dt.$$

Letting $\lambda\to 0^+$, it follows from Theorem 2.2 that $|S(s)x-u(s)|=0$ since u is Lipschitz continuous, which shows the desired result. $\Box$

In general, the semigroup $S(t)$ generated on $\overline{dom\,(A)}$ in Theorem 2.2 in not necessary strongly differentiable. In fact, an example of an m-dissipative A satisfying (5.14) is given, for which the semigroup constructed in Theorem 2.2 is not even weakly differentiable for all $t\ge 0$. Hence, from Theorem 3.1 the corresponding Cauchy problem (6.1) does not have a strong solution. However, we have the following.

Theorem 3.2 Let A be an ω-dissipative subset satisfying (5.14) and $S(t)$, $t\ge 0$ be the semigroup on $\overline{dom\,(A)}$, as constructed in Theorem 2.2. Then, the followings hold.
(1) $u(t)=S(t)x$ on $\overline{dom\,(A)}$ defined in Theorem 2.2 is an integral solution of type ω to the Cauchy problem (6.1).
(2) If $v(t)\in C(0,T;X)$ be an integral of type ω to (6.1), then $|v(t)-u(t)|\le e^{\omega t}\,|v(0)-u(0)|$.
(3) The Cauchy problem (6.1) has a unique solution in $\overline{dom\,(A)}$ in the sense of Definition 2.1.

Proof: A simple modification of the proof of Theorem 2.2 shows that for $x_0\in\overline{dom\,(A)}$

$$S(t)x_0 = \lim_{\lambda\to 0^+}(I-\lambda\, A)^{-[\frac{t}{\lambda}]}x_0$$

exists and defines the semigroup $S(t)$ of nonlinear ω-contractions on $\overline{dom\,(A)}$, i.e.,

$$|S(t)x-S(t)y|\le e^{\omega t}\,|x-y| \quad \text{for } \; t\ge 0 \text{ and } x,y\in\overline{dom\,(A)}.$$

For $x_0\in dom\,(A)$ we define for $\lambda>0$ and $k\ge 1$

$$y_\lambda^k = \lambda^{-1}(J_\lambda^k x_0 - J_\lambda^{k-1}x_0) = A_\lambda J_\lambda^{k-1}x_0 \in AJ_\lambda^k. \tag{6.3}$$

Since A is ω-dissipative, $\langle y_{\lambda,k}-y, J_\lambda^k x_0 - x)\rangle_- \le \omega\,|J_\lambda^k x_0 - x)|$ for $[x,y]\in A$. Since from Lemma 1.1 (4) $\langle y,x\rangle_- - \langle z,x\rangle_+ \le \langle y-z,x\rangle_-$, it follows that

$$\langle y_\lambda^k, J_\lambda^k x_0 - x\rangle_- \le \omega\,|J_\lambda^k x_0 - x)| + \langle y, J_\lambda^k x_0 - x\rangle_+ \tag{6.4}$$

Since from Lemma 1.1 (3) $\langle x+y,x\rangle_- = |x| + \langle y,x\rangle_-$, we have

$$\langle \lambda y_\lambda^k, J_\lambda^k x_0 - x)\rangle_- = |J_\lambda^k x_0 - x| + \langle -(J_\lambda^{k-1}-x), J_\lambda^k x_0 - x\rangle_{\ge} |J_\lambda^k x_0 - x)| - |J_\lambda^{k-1}x_0 - x|.$$

It thus follows from (6.4) that

$$|J_\lambda^k x_0 - x| - |J_\lambda^{k-1} x_0 - x| \leq \lambda\,(\omega\,|J_\lambda^k x_0 - x| + \langle y, J_\lambda^k x_0 - x\rangle_+).$$

Since $J^{[\frac{t}{\lambda}]} = J_\lambda^k$ on $t \in [k\lambda, (k+1)\lambda)$, this inequality can be written as

$$|J_\lambda^k x_0 - x| - |J_\lambda^{k-1} x_0 - x| \leq \int_{k\lambda}^{(k+1)\lambda} (\omega\,|J_\lambda^{[\frac{t}{\lambda}]} x_0 - x| + \langle y, J_\lambda^{[\frac{t}{\lambda}]} x_0 - x\rangle_+\, dt.$$

Hence, summing up this in k from $k = [\frac{\hat{t}}{\lambda}] + 1$ to $[\frac{t}{\lambda}]$ we obtain

$$|J_\lambda^{[\frac{t}{\lambda}]} x_0 - x| - |J_\lambda^{[\frac{\hat{t}}{\lambda}]} x_0 - x| \leq \int_{[\frac{\hat{t}}{\lambda}]\lambda}^{[\frac{t}{\lambda}]\lambda} (\omega\,|J_\lambda^{[\frac{s}{\lambda}]} x_0 - x| + \langle y, J_\lambda^{[\frac{s}{\lambda}]} x_0 - x\rangle_+\, ds.$$

Since $|J_\lambda^{[\frac{s}{\lambda}]} x_0| \leq (1 - \lambda\omega)^{-k}\,|x_0| \leq e^{k\lambda\omega}\,|x_0|$, by Lebesgue dominated convergence theorem and the upper semicontinuity of $\langle \cdot, \cdot \rangle_+$, letting $\lambda \to 0^+$ we obtain

$$|S(t)x_0 - x| - |S(\hat{t})x_0 - x| \leq \int_{\hat{t}}^{t} (\omega\,|S(s)x_0 - x| + \langle y, S(t)x_0 - x\rangle_+)\, ds \tag{6.5}$$

for $x_0 \in dom\,(A)$. Similarly, since $S(t)$ is Lipschitz continuous on $\overline{dom\,(A)}$, again by Lebesgue dominated convergence theorem and the upper semicontinuity of $\langle \cdot, \cdot \rangle_+$, (6.5) holds for all $x_0 \in \overline{dom\,(A)}$.

(2) Let $v(t) \in C(0, T; X)$ be an integral of type ω to (6.1). Since $[J_k^\lambda x_0, y_\lambda^k] \in A$, it follows from (6.2) that

$$|v(t) - J_\lambda^k x_0| - |v(\hat{t}) - J_\lambda^k x_0| \leq \int_{\hat{t}}^{t} (\omega\,|v(s) - J_\lambda^k x_0| + \langle y_\lambda^k, v(s) - J_\lambda^k x_0\rangle_+)\, ds. \tag{6.6}$$

Since $\lambda\, y_\lambda^k = -(v(s) - J_\lambda^k x_0) + (v(s) - J_\lambda^{k-1} x_0)$ and from Lemma 1.1 (3) $\langle -x + y, x\rangle_+ = -|x| + \langle y, x\rangle_+$, we have

$$\langle \lambda y_\lambda^k, v(s) - J_\lambda^k x_0\rangle = -|v(s) - J_\lambda^k x_0| + \langle v(s) - J_\lambda^{k-1} x_0, v(s) - J_\lambda^k x_0\rangle_+ \leq -|v(s) - J_\lambda^k x_0| + |v(s) - J_\lambda^{k-1} x_0|.$$

Thus, from (6.6)

$$(|v(t) - J_\lambda^k x_0| - |v(\hat{t}) - J_\lambda^k x_0|)\lambda \leq \int_{\hat{t}}^{t} (\omega\lambda\,|v(s) - J_\lambda^k x_0| - |v(s) - J_\lambda^k x_0| + |v(s) - J_\lambda^{k-1} x_0|)\, ds$$

Summing up the both sides of this in k from $[\frac{\hat{\tau}}{\lambda}] + 1$ to $[\frac{\tau}{\lambda}]$, we obtain

$$\int_{[\frac{\hat{\tau}}{\lambda}]\lambda}^{[\frac{\tau}{\lambda}]\lambda} (|v(t) - J_\lambda^{[\frac{\sigma}{\lambda}]} x_0| - |v(\hat{t} - J_\lambda^{[\frac{\sigma}{\lambda}]} x_0|\, d\sigma$$

$$\leq \int_{\hat{t}}^{t} (-|v(s) - J_\lambda^{[\frac{\tau}{\lambda}]} x_0| + |v(s) - J_\lambda^{[\frac{\hat{\tau}}{\lambda}]} x_0| + \int_{[\frac{\hat{\tau}}{\lambda}]\lambda}^{[\frac{\tau}{\lambda}]\lambda} \omega\,|v(s) - J_\lambda^{[\frac{\sigma}{\lambda}]} x_0|\, d\sigma)\, ds.$$

Now, by Lebesgue dominated convergence theorem, letting $\lambda \to 0^+$

$$\begin{aligned} &\int_{\hat{\tau}}^{\tau} (|v(t) - u(\sigma)| - |v(\hat{t} - u(\sigma)|)\, d\sigma + \int_{\hat{t}}^{t} (|v(s) - u(\tau)| - |v(s) - u(\hat{\tau})|)\, ds \\ &\qquad \leq \int_{\hat{t}}^{t} \int_{\hat{\tau}}^{\tau} \omega\,|v(s) - u(\sigma)|\, d\sigma\ ds. \end{aligned} \tag{6.7}$$

For $h>0$ we define F_h by

$$F_h(t)=h^{-2}\int_t^{t+h}\int_t^{t+h}|v(s)-u(\sigma)|\,d\sigma\,ds.$$

Then from (6.7) we have $\dfrac{d}{dt}F_h(t)\leq \omega\,F_h(t)$ and thus $F_h(t)\leq e^{\omega t}F_h(0)$. Since u, v are continuous we obtain the desired estimate by letting $h\to 0^+$. □

Lemma 3.3 Let A be an ω-dissipative subset satisfying (5.14) and $S(t)$, $t\geq 0$ be the semigroup on $\overline{dom\,(A)}$, as constructed in Theorem 2.2. Then, for $x_0\in\overline{dom\,(A)}$ and $[x,y]\in A$

$$|S(t)x_0-x|^2-|S(\hat{t})x_0-x|^2\leq 2\int_{\hat{t}}^t(\omega\,|S(s)x_0-x|^2+\langle y,S(s)x_0-x\rangle_s\,ds \tag{6.8}$$

and for every $f\in F(x_0-x)$

$$\limsup_{t\to 0^+}\,Re\,\langle\frac{S(t)x_0-x_0}{t},f\rangle\leq\omega\,|x_0-x|^2+\langle y,x_0-x\rangle_s. \tag{6.9}$$

Proof: Let y_k^λ be defined by (6.3). Since A is ω-dissipative, there exists $f\in F(J_\lambda^k x_0-x)$ such that $Re\,(y_\lambda^k-y,f\rangle\leq\omega\,|J_\lambda^k x_0-x|^2$. Since

$$Re\,\langle y_\lambda^k,f\rangle=\lambda^{-1}\,Re\,\langle J_\lambda^k x_0-x-(J_\lambda^{k-1}x_0-x),f\rangle$$

$$\geq\lambda^{-1}(|J_\lambda^k x_0-x|^2-|J_\lambda^{k-1}x_0-x||J_\lambda^k x_0-x|)\geq(2\lambda)^{-1}(|J_\lambda^k x_0-x|^2-|J_\lambda^{k-1}x_0-x|^2),$$

we have from Theorem 1.4

$$|J_\lambda^k x_0-x|^2-|J_\lambda^{k-1}x_0-x|^2\leq 2\lambda\,Re\,\langle y_\lambda^k,f\rangle\leq 2\lambda\,(\omega\,|J_\lambda^k x_0-x|^2+\langle y,J_\lambda^k x_0-x\rangle_s.$$

Since $J_\lambda^{[\frac{t}{\lambda}]}x_0=J_\lambda^k x_0$ on $[k\lambda,(k+1)\lambda)$, this can be written as

$$|J_\lambda^k x_0-x|^2-|J_\lambda^{k-1}x_0-x|^2\leq\int_{k\lambda}^{(k+1)\lambda}(\omega\,|J_\lambda^{[\frac{t}{\lambda}}x_0-x|^2+\langle y,J_\lambda^{[\frac{t}{\lambda}]}x_0-x\rangle_s)\,dt.$$

Hence,

$$|J_\lambda^{[\frac{t}{\lambda}]}x_0-x|^2-|J_\lambda^{[\frac{\hat{t}}{\lambda}]}x_0-x|^2\leq 2\lambda\int_{[\frac{\hat{t}}{\lambda}]\lambda}^{[\frac{t}{\lambda}]\lambda}(\omega\,|J_\lambda^{[\frac{s}{\lambda}]}x_0-x|^2+\langle y,J_\lambda^{[\frac{s}{\lambda}]}x_0-x\rangle_s)\,ds.$$

Since $|J_\lambda^{[\frac{s}{\lambda}]}x_0|\leq(1-\lambda\omega)^{-k}\,|x_0|\leq e^{k\lambda\omega}\,|x_0|$, by Lebesgue dominated convergence theorem and the upper semicontinuity of $\langle\cdot,\cdot\rangle_s$, letting $\lambda\to 0^+$ we obtain (6.8).

Next, we show (6.9). For any given $f\in F(x_0-x)$ as shown above

$$2Re\,\langle S(t)x_0-x_0,f\rangle\leq|S(t)x_0-x|^2-|x_0-x|^2.$$

Thus, from (6.8)

$$Re\,\langle S(t)x_0-x_0,f\rangle\leq\int_0^t(\omega\,|S(s)x_0-x|^2+\langle y,S(s)x_0-x\rangle_s)\,ds$$

Since $s \to S(s)x_0$ is continuous, by the upper semicontinuity of $\langle \cdot, \cdot \rangle_s$, we have (6.9). $\square$

Theorem 3.4 Assume that A is a close dissipative subset of $X \times X$ and satisfies the range condition (5.14) and let $S(t),\ t \geq 0$ be the semigroup on $\overline{dom\,(A)}$, defined in Theorem 2.2. Then, if $S(t)x$ is strongly differentiable at $t_0 > 0$ then

$$S(t_0)x \in dom\,(A) \quad \text{and} \quad \frac{d}{dt}S(t)x\,|_{t=t_0} \in AS(t)x,$$

and moreover

$$S(t_0)x \in dom\,(A^0) \quad \text{and} \quad \frac{d}{dt}S(t)x\,\Big|_{t=t_0} = A^0 S(t_0)x.$$

Proof: Let $\frac{d}{dt}S(t)x\,|_{t=t_0} = y$. Then $S(t_0-\lambda)x - (S(t_0)x - \lambda\, y) = o(\lambda)$, where $\frac{|o(\lambda)|}{\lambda} \to 0$ as $\lambda \to 0^+$. Since $S(t_0-\lambda)x \in \overline{dom\,(A)}$, there exists a $[x_\lambda, y_\lambda] \in A$ such that $S(t_0-\lambda)x = x_\lambda - \lambda\, y_\lambda$ and

$$\lambda\,(y - y_\lambda) = S(t_0)x - x_\lambda + o(\lambda). \tag{6.10}$$

If we let $x = x_\lambda$, $y = y_\lambda$ and $x_0 = S(t_0)x$ in (6.9), then we obtain

$$Re\,\langle y, f\rangle \leq \omega\,|S(t_0)x - x_\lambda|^2 + \langle y_\lambda, S(t_0)x - x_\lambda\rangle_s.$$

for all $f \in (S(t_0)x - x_\lambda)$. It follows from Lemma 1.3 that there exists a $g \in F(S(t_0)x - x_\lambda)$ such that $\langle y_\lambda, S(t)x - x_\lambda\rangle_s = Re\,\langle y_\lambda, g\rangle$ and thus

$$Re\,\langle y - y_\lambda, g\rangle \leq \omega\,|S(t_0)x - x_\lambda|^2.$$

From (6.10)

$$\lambda^{-1}|S(t_0)x - x_\lambda|^2 \leq \frac{|o(\lambda)|}{\lambda}\,|S(t_0)x - x_\lambda| + \omega\,|S(t_0)x - x_\lambda|^2$$

and thus

$$(1-\lambda\omega)\left|\frac{S(t_0)x - x_\lambda}{\lambda}\right| \to 0 \quad \text{as} \quad \lambda \to 0^+.$$

Combining this with (6.10), we obtain

$$x_\lambda \to S(t_0)x \quad \text{and} \quad y_\lambda \to y$$

as $\lambda \to 0^+$. Since A is closed, it follows that $[S(t_0)x, y] \in A$, which shows the first assertion.

Next, from Theorem 2.2

$$|S(t_0+\lambda)x - S(t_0)x| \leq \lambda\,\|AS(t_0)x\| \quad \text{for} \quad \lambda > 0$$

This implies that $|y| \leq \|AS(t_0)x\|$. Since $y \in AS(t_0)x$, it follows that $S(t_0)x \in dom\,(A^0)$ and $y \in A^0 S(t_0)x$. $\square$

Theorem 3.5 Let A be a dissipative subset of $X \times X$ satisfying the range condition (5.14) and $S(t),\ t \geq 0$ be the semigroup on $\overline{dom\,(A)}$, defined in Theorem 2.2. Then the followings hold.

(1) For ever $x \in \overline{dom\,(A)}$

$$\lim_{\lambda \to 0^+} |A_\lambda x| = \liminf_{t \to 0^+} |S(t)x - x|.$$

(2) Let $x \in \overline{dom\,(A)}$. Then, $\lim_{\lambda\to 0^+} |A_\lambda x| < \infty$ if and only if there exists a sequence $\{x_n\}$ in $dom\,(A)$ such that $\lim_{n\to\infty} x_n = x$ and $\sup_n \|Ax_n\| < \infty$.

Proof: Let $x \in \overline{dom\,(A)}$. From Theorem 1.4 $|A_\lambda x|$ is monotone decreasing and $\lim |A_\lambda x|$ exists (including ∞). Since $|J_\lambda^{[\frac{t}{\lambda}]} - x| \leq t\,|A_\lambda x|$, it follows from Theorem 2.2 that

$$|S(t)x - x| \leq t \lim_{\lambda\to 0^+} |A_\lambda x|$$

thus

$$\liminf_{t\to 0^+} \frac{1}{t}\,|S(t)x - x| \leq \lim_{\lambda\to 0^+} |A_\lambda x|.$$

Conversely, from Lemma 3.3

$$-\liminf_{t\to 0^+} \frac{1}{t}\,|S(t)x - x||x-u| \leq \langle v, x-u\rangle_s$$

for $[u, v] \in A$. We set $u = J_\lambda x$ and $v = A_\lambda x$. Since $x - u = -\lambda\, A_\lambda x$,

$$-\liminf_{t\to 0^+} \frac{1}{t}\,|S(t)x - x|\lambda|A_\lambda x| \leq -\lambda\,|A_\lambda x|^2$$

which implies

$$\liminf_{t\to 0^+} \frac{1}{t}\,|S(t)x - x| \geq |A_\lambda x|.$$

Theorem 3.7 Let A be a dissipative set of $X \times X$ satisfying the range condition (5.14) and let $S(t),\ t \geq 0$ be the semigroup defined on $\overline{dom\,(A)}$ in Theorem 2.2.

(1) If $x \in dom\,(A)$ and $S(t)x$ is differentiable a.e., $t > 0$, then $u(t) = S(t)x,\ t \geq 0$ is a unique strong solution of the Cauchy problem (6.1).

(2) If X is reflexive. Then, if $x \in dom\,(A)$ then $u(t) = S(t)x,\ t \geq 0$ is a unique strong solution of the Cauchy problem (6.1).

Proof: The assertion (1) follows from Theorems 3.1 and 3.5. If X is reflexive, then since an X-valued absolute continuous function is a.e. strongly differentiable, (2) follows from (1). □

6.0.3 Infinitesimal generator

Definition 4.1 Let X_0 be a subset of a Banach space X and $S(t),\ t \geq 0$ be a semigroup of nonlinear contractions on X_0. Set $A_h = h^{-1}(T(h) - I)$ for $h > 0$ and define the strong and weak infinitesimal generators A_0 and A_w by

$$A_0 x = \lim_{h\to 0^+} A_h x \quad \text{with } dom\,(A_0) = \{x \in X_0 : \lim_{h\to 0^+} A_h x \text{ exists}\}$$

$$A_0 x = w - \lim_{h\to 0^+} A_h x \quad \text{with } dom\,(A_0) = \{x \in X_0 : w - \lim_{h\to 0^+} A_h x \text{ exists}\}, \tag{6.11}$$

respectively. We define the set $\hat{D}$ by

$$\hat{D} = \{x \in X_0 : \liminf_{h\to 0} |A_h x| < \infty\} \tag{6.12}$$

Theorem 4.1 Let $S(t),\ t \geq 0$ be a semigroup of nonlinear contractions defined on a closed subset X_0 of X. Then the followings hold.
(1) $\langle A_w x_1 - A_w x_2, x^* \rangle$ for all $x_1, x_2 \in dom\,(A_w)$ and $x^* \in F(x_1 - x_2)$. In particular, A_0 and A_w are dissipative.
(2) If X is reflexive, then $\overline{dom\,(A_0)} = \overline{dom\,(A_w)} = \overline{\hat{D}}$.
(3) If X is reflexive and strictly convex, then $dom(A_w) = \hat{D}$. In addition, if X is uniformly convex, then $dom\,(A_w) = dom\,(A_0) = \hat{D}$ and $A_w = A_0$.

Proof: (1) For $x_1, x_2 \in X_0$ and $x^* \in F(x_1 - x_2)$ we have

$$\langle A_h x_1 - A_h x_2, x^* \rangle = h^{-1}(\langle S(h)x_1 - S(h)x_2, x^* \rangle - |x_1 - x_2|^2)$$

$$\leq h^{-1}(|S(h)x_1 - S(h)x_2||x_1 - x_2| - |x_1 - x_2|^2) \leq 0.$$

Letting $h \to 0^+$, we obtain the desired inequality.
(2) Obviously, $dom\,(A_0) \subset dom\,(A_w) \subset \hat{D}$. Let $x \in \hat{D}$. It suffices to show that $x \in \overline{dom\,(A_0)}$. We can show that $t \to S(t)x$ is Lipschitz continuous. In fact, there exists a monotonically decreasing sequence $\{t_k\}$ of positive numbers and $L > 0$ such that $t_k \to 0$ as $k \to \infty$ and $|S(t_k)x - x| \leq L\,t_k$. Let $h > 0$ and n_k be a nonnegative integer such that $0 \leq h - n_k t_k < t_k$. Then we have

$$|S(t+h)x - S(t)x| \leq |S(t)x - x| = |S(h - n_k t_k + n_k t_k)x - x|$$

$$\leq |S(h - n_k t_k)x - x| + L\,n_k t_k \leq |S(h - n_k t_k)x - x| + L\,h$$

By the strong continuity of $S(t)x$ at $t = 0$, letting $k \to \infty$, we obtain $|S(t+h)x - S(t)x| \leq L\,h$. Now, since X is reflexive this implies that $S(t)x$ is a.e. differentiable on $(0, \infty)$. But since $S(t)x \in dom\,(A_0)$ whenever $\frac{d}{dt}S(t)x$ exists, $S(t)x \in dom\,(A_0)$ a.e. $t > 0$. Thus, since $|S(t)x - x| \to 0$ as $t \to 0^+$, it follows that $x \in \overline{dom\,(A_0)}$.
(3) Assume that X is reflexive and strictly convex. Let $x_0 \in \hat{D}$ and Y be the set of all weak cluster points of $t^{-1}(S(t)x_0 - x_0)$ as $t \to 0^+$. Let $\tilde{A}$ be a subset of $X \times X$ defined by

$$\tilde{A} = A_0 \cup [x_0, \overline{co\,Y}] \quad \text{and} \quad dom\,(\tilde{A}) = dom\,(A_0) \cup \{x_0\}$$

where $\overline{co\,Y}$ denotes the closure of the convex hull of Y. Note that from (1)

$$\langle A_0 x_1 - A_0 x_2, x^* \rangle \leq 0 \quad \text{for all } \ x_1, x_2 \in dom\,(A_0) \text{ and } x^* \in F(x_1 - x_2)$$

and for every $y \in Y$

$$\langle A_0 x_1 - y, x^* \rangle \leq 0 \quad \text{for all } \ x_1 \in dom\,(A_0) \text{ and } x^* \in F(x_1 - x_0).$$

This implies that $\tilde{A}$ is a dissipative subset of $X \times X$. But, since X is reflexive, $t \to S(t)x_0$ is a.e. differentiable and

$$\frac{d}{dt}S(t)x_0 = A_0 S(t)x_0 \in \tilde{A}S(t)x_0, \quad a.e.,\ t > 0.$$

It follows from the dissipativity of $\tilde{A}$ that

$$\langle \frac{d}{dt}(S(t)x_0 - x_0), x^* \rangle \leq \langle y, x^* \rangle, \quad a.e,\ t > 0 \text{ and } y \in \tilde{A}x_0 \tag{6.13}$$

for $x^* \in F(S(t)x_0 - x_0)$. Note that for $h > 0$

$$\langle h^{-1}(S(t+h)x_0 - S(t)x_0), x^*\rangle \le h^{-1}(|S(t+h)x_0 - x_0| - |S(t)x_0 - x_0|)|x^*| \quad \text{for } x^* \in F(S(t)x_0 - x_0).$$

Letting $h \to 0^+$, we have

$$\langle \frac{d}{dt}(S(t)x_0 - x_0), x^*\rangle \le |S(t)x_0 - x_0| \frac{d}{dt}|S(t)x_0 - x_0|, \quad a.e.\ t > 0$$

The converse inequality follows much similarly. Thus, we have

$$|S(t)x_0 - x_0|\frac{d}{dt}|S(t)x_0 - x_0| = \langle \frac{d}{dt}(S(t)x_0 - x_0), x^*\rangle \quad \text{for } \ x^* \in F(S(t)x_0 - x_0). \tag{6.14}$$

It follows from (6.13)–(6.14) that $\frac{d}{dt}|S(t)x_0 - x_0| \le |y|$ for $y \in Y$ and a.e. $t > 0$ and thus

$$|S(t)x_0 - x_0| \le t\,\|\tilde{A}x_0\| \quad \text{for all } t > 0. \tag{6.15}$$

Note that $\tilde{A}x_0 = \overline{co\, Y}$ is a closed convex subset of X. Since X is reflexive and strictly convex, there exists a unique element $y_0 \in \tilde{A}x_0$ such that $|y_0| = \|\tilde{A}x_0\|$. Hence, (6.15) implies that $\overline{co\, Y} = y_0 = A_w x_0$ and therefore $x_0 \in dom\,(A_w)$.

Next, we assume that X is uniformly convex and let $x_0 \in dom(A_w) = \hat{D}$. Then

$$w - \lim \frac{S(t)x_0 - x_0}{t} = y_0 \quad \text{as } \ t \to 0^+.$$

From (6.15)

$$|t^{-1}(S(t)x_0 - x_0)| \le |y_0|, \quad a.e.\ t > 0.$$

Since X is uniformly convex, these imply that

$$\lim \frac{S(t)x_0 - x_0}{t} = y_0 \quad \text{as } \ t \to 0^+.$$

which completes the proof. □

Theorem 4.2 Let X and X^* be uniformly convex Banach spaces. Let $S(t)$, $t \ge 0$ be the semigroup of nonlinear contractions on a closed subset X_0 and A_0 be the infinitesimal generator of $S(t)$. If $x \in dom\,(A_0)$, then
(*i*) $S(t)x \in dom\,(A_0)$ for all $t \ge 0$ and the function $t \to A_0 S(t)x$ is right continuous on $[0, \infty)$.
(*ii*) $S(t)x$ has a right derivative $\frac{d^+}{dt}S(t)x$ for $t \ge 0$ and $\frac{d^+}{dt}S(t)x = A_0 S(t)x$, $t \ge 0$.
(*iii*) $\frac{d}{dt}S(t)x$ exists and is continuous except a countable number of values $t \ge 0$.

Proof: $(i) - (ii)$ Let $x \in dom\,(A_0)$. By Theorem 4.1, $dom\,(A_0) = \hat{D}$ and thus $S(t)x \in dom\,(A_0)$ and

$$\frac{d^+}{dt}S(t)x = A_0 S(t)x \quad \text{for } \ t \ge 0. \tag{6.16}$$

Moreover, $t \to S(t)x$ a.e. differentiable and $\frac{d}{dt}S(t)x = A_0 S(t)x$ a.e. $t > 0$. We next prove that $A_0 S(t)x$ is right continuous. For $h > 0$

$$\frac{d}{dt}(S(t+h)x - S(t)x) = A_0 S(t+h)x - A_0 S(t)x, \quad a.e.\ t > 0.$$

From (6.14)

$$|S(t+h)x - S(t)x|\,\frac{d}{dt}|S(t+h)x - S(t)x| = \langle A_0S(t+h)x - A_0S(t)x, x^*\rangle \le 0$$

for all $x^* \in F(S(t+h)x - S(t)x)$, since A_0 is dissipative. Integrating this over $[s,t]$, we obtain

$$|S(t+h)x - S(t)x| \le |S(s+h)x - S(s)x| \quad \text{for } \ 0 \le s \le t$$

and therefore

$$|\frac{d^+}{dt}S(t)x| \le |\frac{d^+}{ds}S(s)x|.$$

Hence $t \to |A_0S(t)x|$ is monotonically non-increasing function and thus it is right continuous. Let $t_0 \ge 0$ and let $\{t_k\}$ be a decreasing sequence of positive numbers such that $t_k \to t_0$. Without loss of generality, we may assume that $w\!-\!\lim_{k\to\infty} A_0S(t_k) = y_0$. The right continuity of $|A_0S(t)x|$ at $t = t_0$, thus implies that

$$|y_0| \le |A_0S(t_0)x| \tag{6.17}$$

since norm is weakly lower semicontinuous. Let $\tilde{A}_0$ be the maximal dissipative extension of A_0. It then follows from Theorem 1.9 that $\tilde{A}$ is demiclosed and thus $y_0 \in \tilde{A}S(t)x$. On the other hand, for $x \in dom\,(A_0)$ and $y \in \tilde{A}x$, we have

$$\langle\frac{d}{dt}(S(t)x - x), x^*\rangle \le \langle y, x^*\rangle \quad \text{for all } \ x^* \in F(S(t)x - x)$$

a.e. $t > 0$, since $\tilde{A}$ is dissipative and $\frac{d}{dt}S(t)x = A_0S(t)x \in \tilde{A}S(t)x$ a.e. $t > 0$. From (6.14) we have

$$t^{-1}|S(t)x - x| \le |\tilde{A}x| = \|\tilde{A}^0x\| \quad \text{for } \ t \ge 0$$

where $\tilde{A}^0$ is the minimal section of $\tilde{A}$. Hence $A_0x = \tilde{A}^0x$. It thus follows from (6.17) that $y_0 = A_0S(t_0)x$ and $\lim_{k\to\infty} A_0S(t_k)x = y_0$ since X is uniformly convex. Thus, we have proved the right continuity of $A_0S(t)x$ for $t \ge 0$.
(iii) Integrating (6.16) over $[t, t+h]$, we have

$$S(t+h)x - S(t)x = \int_t^{t+h} A_0S(s)x\,ds$$

for $t,\ h \ge 0$. Hence it suffices to prove that the function $t \to A_0S(t)x$ is continuous except a cuntable number of $t > 0$. Using the same arguments as above, we can show that if $|A_0S(t)x|$ is continuous at $t = t_0$, then $A_0S(t)x$ is continuous at $t = t_0$. But since $|A_0S(t)x|$ is monotone non-increasing, it follows that it has at most countably many discontinuities, which completes the proof. □

Theorem 4.3 Let X and X^* be uniformly convex Banach spaces. If A be m-dissipative, then A is demiclosed, $\overline{dom\,(A)}$ is a closed convex set and A^0 is single-valued operator with $dom\,(A^0) = dom\,(A)$. Moreover, A^0 is the infinitesimal generator of a semigroup of contractions on $\overline{dom\,(A)}$.

Proof: It follows from Theorem 1.9 that A is demiclosed. The second assertion follows from Theorem 1.12. Also, from Theorem 3.4

$$\frac{d}{dt}S(t)x = A^0S(t)x, \quad a.e.\ t > 0$$

and

$$|S(t)x - x| \leq t\,|A^0 x|, \quad t > 0 \tag{6.18}$$

for $x \in dom\,(A)$. Let A_0 be the infinitesimal generator of the semigroup $S(t)$, $t \geq 0$ generated by A defined in Theorem 2.2. Then, (6.18) implies that by Theorem 4.1 $x \in dom\,(A_0)$ and by Theorem 4.2 $\frac{d^+}{dt} S(t)x = A_0 S(t)x$ and $A_0 S(t)x$ is right continuous in t. Since A is closed,

$$A_0 x = \lim_{t \to 0^+} A_0 S(t)x \in Ax.$$

Hence, (6.17) implies that $A_0 x = A^0 x$.

When X is a Hilbert space we have the nonlinear version of Hille-Yosida theorem as follows.

Theorem 4.4 Let H be a Hilbert space. Then,
(1) The infinitesimal generator A_0 of a semigroup of contractions $S(t)$, $t \geq 0$ on a closed convex set X_0 has a dense domain in X_0 and there exists a unique maximal dissipative operator A such that $A^0 = A_0$.

Conversely,
(2) If A_0 is a maximal dissipative operator, then $\overline{dom\,(A)}$ is a closed convex set and A^0 is the infinitesimal generator of contractions on $\overline{dom\,(A)}$.

Proof: (2) Since from Theorem 1.7 the maximal dissipative operator in a Hilbert space is m-dissipative, (2) follows from Theorem 4.3.
Example (Nonlinear Diffusion)

$$u_t = Au = \Delta\gamma(u) + \nabla \cdot (\vec{f}(u))$$

on $X = L^1$. Assume $\gamma : R \to R$ is strictly monotone. Thus, $sign(x-y) = sign(\gamma(x) - \gamma(y))$

$$(Au_1 - Au_2, \rho_\epsilon(\gamma(u_1) - \gamma(u_2))) = -(\nabla(\gamma(u_1) - \gamma(u_2))\rho_\epsilon')\nabla(\gamma(u_1) - \gamma(u_2))) - (\rho_\epsilon'(u_1 - u_2))(f(u_1) - f(u_2)), \nabla(u_1 -$$

6.1 Fixed Point Theorems

Banach Fixed Point Theorem. Let (X, d) be a non-empty complete metric space with a contraction mapping $T : X \to X$. Then T admits a unique fixed-point x^* in X (i.e. $T(x*) = x*$). Moreover, x^* is a fixed point of the fixed point iterate $x_n = T(x_{n-1})$ with arbitrary initial condition x_0.

Proof:

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1}) \leq \rho\, d(x_n, x_{n-1}.$$

By the induction

$$d(x_{n+1}, x_n) \leq \rho^n\, d(x_1, x_0)$$

and

$$d(x_m, x_n) \leq \frac{\rho^{n-1}}{1-\rho}\, d(x_1, x_0) \to 0$$

as $m \geq n \to \infty$. Thus, $\{x_n\}$ is a Cauchy sequence and $\{x_n\}$ converges a fixed point x^* of T. Suppose x_1^*, x_2^* are two fixed point. Then,

$$d(x_1^*, x_2^*) = d(T(x_1^*), T(x_2^*) \leq \rho\, d(x_1^*, x_2^*)$$

Since $\rho < 1$, $d(x_1^*, x_2^*) = 0$ and thus $x_1^* = x_2^*$. $\square$

Brouwer fixed point theorem: is a fundamental result in topology which proves the existence of fixed points for continuous functions defined on compact, convex subsets of Euclidean spaces. Kakutani's theorem extends this to set-valued functions.

Kakutani's theorem states: Let S be a non-empty, compact and convex subset of some Euclidean space R^n. Let $T : S \to S$ be a set-valued function on S with a closed graph and the property that $T(x)$ is non-empty and convex for all $x \in S$. Then T has a fixed point.

The Schauder fixed point theorem is an extension of the Brouwer fixed point theorem to topological vector spaces.

Schauder fixed point theorem Let X be a locally convex topological vector space, and let $K \subset X$ be a compact, and convex set. Then given any continuous mapping $T : K \to K$ has a fixed point.

Given $\epsilon > 0$ there exists the family of open sets $\{B_\epsilon?(x) : x \in K\}$ is open covering of K Since K is compact. there exists a finite open subcover, i.e. there exist finite many points $\{x_k\}$ in K such that

$$K = \bigcup_{k=1}^{n} B_\epsilon(x_k)$$

Define the functions $\{g_k\}$ by $g_k(x) = \max(0, \epsilon - |x - x_k|)$. It is clear that each g_k is continuous, $g_k(x) \geq 0$ and $\sum_k g_k(x) > 1$ for all $x \in K$. Thus, we can define a function on K by

$$g?(x) = \frac{\sum_k g_k(x)\, x_k}{\sum_k g_k(x)}$$

and g is a continuous function from K to the convex hull K_0 of of $\{x_k\}$. It is easily shown that $|g?(x) - x| \leq \epsilon$ for $x \in K$. Then A maps K_0 to K_0Since K_0 is compact convex subset of a finite dimensional vector space, we can apply the Brouwer fixed point theorem to assure the existence of $z_\epsilon \in K_0$ such that $x_\epsilon = A(x_\epsilon)$. We have the estimate

$$|z_\epsilon - T(z_\epsilon)| = |g(T(z_\epsilon)) - T(z_\epsilon)| \leq \epsilon.$$

Since K is compact, the exits a subsequence $\epsilon_n \to 0$ such that $x_{\epsilon_n} \to x_0 \in K$ and thus

$$|x_{\epsilon_n} - T(x_{\epsilon_n})| \leq |g(T(x_{\epsilon_n})) - T(x_{\epsilon_n})| \leq \epsilon_n.$$

Since T is continuous, $T(x_0) = x_0$. $\square$

Schaefer's fixed point theorem: Let T be a continuous and compact mapping of a Banach space X into itself, such that the set

$$\{x \in X : x = \lambda T x \text{ for some } 0 \leq \lambda \leq 1\}$$

is bounded. Then T has a fixed point.

Proof: Let $|x| \leq M$ for $\{x \in X : x = \lambda T x$ for some $0 \leq \lambda \leq 1\}$. Define

$$\tilde{T}(x) = \frac{M T(x)}{\max(M, |T(x)|)}$$

Note that $\tilde{T} : B(0, M) \to B(0, M)$.. Let K be the closed convex hull go $\tilde{T}(B(0, M))$. Since T is compact, K is a compact convex subset of X. Since $\tilde{T}K \to K$, it follows from Schauder's fixed point theorem that there exits $x \in K$ such that $x = \tilde{T}(x)$. We now claim that x is a fixed point of T. If not, we should have $|T(x)| > M$ and

$$x = \lambda T(u) \text{ for } \lambda = \frac{M}{|T(x)|} < 1$$

But, since $|x| = |\tilde{T}(x)| \leq M$, which is a contradiction. $\square$

The advantage of Schaefer's theorem over Schauder's for applications is that it is not necessary to determine an explicit convex, compact set K.

Krasnoselskii theorem The sum of two operators $T = A + B$ has a fixed point in a nonempty closed convex subset C of a real Banach space X under conditions such that $T(C) \subset C$, A is continuous on C, $A(C)$ is a relatively compact subset of X, and B is a strict contraction on C.

Note that the proof of Kransnoselskiis fixed point theorem combines the Banach contraction theorem and Schauders fixed point theorem, i.e., $y \in C$, $Ay + Bx = x$ has the fixed point map $x = \psi(y)$ and use Schauders fixed point theorem for ψ.
Kakutani-Glicksberg-Fan theorem Let S be a non-empty, compact and convex subset of a locally convex topological vector space X. Let $T : S \to S$ be a Kakutani map, i.e., it is upper hemicontinuous, i.e., if for every open set $W \subset X$, the set $\{x \in X : T(x) \in W\}$ is open in X and $T(x)$ is non-empty, compact and convex for all $x \in X$. Then T has a fixed point.

The corresponding result for single-valued functions is the Tychonoff fixed-point theorem.

7 Convex Analysis and Duality

This chapter is organized as follows. We present the basic convex analysis and duality theory and subdifferntial and monotone operator theory in Banach spaces.

7.1 Convex Functional and Subdifferential

Definition (Convex Functional) (1) A proper convex functional on a Banach space X is a function φ from X to $(-\infty, \infty]$, not identically $+\infty$ such that

$$\varphi((1-\lambda)\, x_1 + \lambda\, x_2) \leq (1-\lambda)\, \varphi(x_1) + \lambda\, \varphi(x_2)$$

for all $x_1,\ x_2 \in X$ and $0 \leq \lambda \leq 1$.
(2) A functional $\varphi : X \to R$ is said to be lower-semicontinuous if

$$\varphi(x) \leq \liminf_{y \to x}\ \varphi(y) \quad \text{for all } \ x \in X.$$

(3) A functional $\varphi : X \to R$ is said to be weakly lower-semicontinuous if

$$\varphi(x) \leq \liminf_{n \to \infty}\ \varphi(x_n)$$

for all weakly convergent sequence $\{x_n\}$ to x.

(4) The subset $D(\varphi) = \{x \in X; \varphi(x) \leq \infty\}$ of X is called the domain of φ.
(5) The epigraph of φ is defined by $epi(\varphi) = \{(x, c) \in X \times R : \varphi(x) \leq c\}$.

Lemma 3 A convex functional φ is lower-semicontinuous if and only if it is weakly lower-semicontinuous on X.

Proof: Since the level set $\{x \in X : \varphi(x) \leq c\}$ is a closed convex subset if φ is lower-semicontinuous. Thus, the claim follows the fact that a convex subset of X is closed if and only if it is weakly closed.

Lemma 4 If φ be a proper lower-semicontinuous, convex functional on X, then φ is bounded below by an affine functional, i.e., there exist $x^* \in X^*$ and $c \in R$ such that

$$\varphi(x) \geq \langle x, x^* \rangle + \beta, \quad x \in X.$$

Proof: Let $x_0 \in X$ and $\beta \in R$ be such that $\varphi(x_0) > c$. Since φ is lower-semicontinuous on X, there exists an open neighborhood $V(x_0)$ of X_0 such that $\varphi(x) > c$ for all $x \in V(x_0)$. Since the ephigraph $epi(\varphi)$ is a closed convex subset of the product space $X \times R$. It follows from the separation theorem for convex sets that there exists a closed hyperplane $H \subset X \times R$;

$$H = \{(x, r) \in X \times R : \langle x, x_0^* \rangle + r = \alpha\} \quad \text{with } \ x_0^* \in X^*, \ \alpha \in R,$$

that separates $epi(\varphi)$ and $V(x_0) \times (-\infty, c)$. Since $\{x_0\} \times (-\infty, c) \subset \{(x, r) \in X \times R : \langle x, x_0^* \rangle + r < \alpha\}$ it follows that

$$\langle x, x_0^* \rangle + r > \alpha \quad \text{for all } \ (x, c) \in epi(\varphi)$$

which yields the desired estimate.

Definition (Subdifferential) Given a proper convex functional φ on a Banach space X the subdifferential of $\partial\varphi(x)$ is a subset in X^*, defined by

$$\partial\varphi(x) = \{x^* \in X^* : \varphi(y) - \varphi(x) \geq \langle y - x, x^* \rangle \ \text{ for all } \ y \in X\}.$$

Since for $x_1^* \in \partial\varphi(x_1)$ and $x_2^* \in \partial\varphi(x_2)$,

$$\varphi(x_1) - \varphi(x_2) \leq \langle x_1 - x_2, x_2^* \rangle$$

$$\varphi(x_2) - \varphi(x_1) \leq \langle x_2 - x_1, x_1^* \rangle$$

it follows that $\langle x_1 - x_2, x_1^* - x_2^* \rangle \geq 0$. Hence $\partial\varphi$ is a monotone operator from X into X^*.

Example 1 Let φ be Gateaux differentiable at x. i.e., there exists $w^* \in X^*$ such that

$$\lim_{t \to 0^+} \frac{\varphi(x + t\,v) - \varphi(x)}{t} = \langle h, w^* \rangle \quad \text{for all } \ h \in X$$

and w^* is the Gateaux differential of φ at x and is denoted by $\varphi'(x)$. If φ is convex, then φ is subdifferentiable at x and $\partial\varphi(x) = \{\varphi'(x)\}$. Indeed, for $v = y - x$

$$\frac{\varphi(x + t\,(y - x)) - \varphi(x)}{t} \leq \varphi(y) - \varphi(x), \quad 0 < t < 1$$

Letting $t \to 0^+$ we have

$$\varphi(y) - \varphi(x) \geq \langle y - x, \varphi'(x) \rangle \quad \text{for all } \ y \in X,$$

and thus $\varphi'(x) \in \partial\varphi(x)$. On the other hand if $w^* \in \partial\varphi(x)$ we have for $y \in X$ and $t > 0$

$$\frac{\varphi(x + t\,y) - \varphi(x)}{t} \geq \langle y, w^* \rangle.$$

Taking limit $t \to 0^+$, we obtain

$$\langle y, \varphi'(x) - w^* \rangle \geq 0 \quad \text{for all } \ y \in X.$$

This implies $w^* = \varphi'(x)$.

Example 2 If $\varphi(x) = \frac{1}{2}\,|x|^2$ then we will show that $\partial\varphi(x) = F(x)$, the duality mapping. In fact, if $x^* \in F(x)$, then

$$\langle x - y, x^* \rangle = |x|^2 - \langle y, x^* \rangle \geq \frac{1}{2}\,(|x|^2 - |y|^2) \quad \text{for all } \ y \in X.$$

Thus $x^* \in \partial\varphi(x)$. Conversely, if $x^* \in \partial\varphi(x)$ then

$$|x + t\,y|^2 \geq |x|^2 + 2t\,\langle y - x, x^* \rangle \quad \text{for all } \ y \in X \ \text{ and } \ t > 0.$$

It can be shown that this implies $|x^*| = |x|$ and $\langle x, x^* \rangle = |x|^2$. Hence $x^* \in F(x)$.

Example 3 Let K be a closed convex subset of X and I_K be the indicator function of K, i.e.,

$$I_K(x) = \begin{cases} 0 & \text{if } x \in K \\ \infty & \text{otherwise.} \end{cases}$$

Obviously, I_K is convex and lower-semicontinuous on X. By definition we have for $x \in K$

$$\partial I_K(x) = \{x^* \in X^* : \langle x - y, x^* \rangle \geq 0 \ \text{ for all } \ y \in K\}$$

Thus $D(I_K) = D(\partial I_K) = K$ and $\partial_K(x) = \{0\}$ for each interior point of K. Moreover, if x lies on the boundary of K, then $\partial I_K(x)$ coincides with the cone of normals to K at x.

Definition (Conjugate function) Given a proper functional φ on a Banach space X, the conjugate functional on X^* is defined by

$$\varphi^*(x^*) = \sup_y \{\langle x^*, x \rangle - \varphi(x)\}.$$

Theorem 2.1 The conjugate function φ^* is convex lower semi-continuous and proper.

Proof: For $0 < t < 1$ and $x_1^*,\ x_2^* \in X^*$

$$\varphi^*(t\,x_1^* + (1-t)\,x_2^*) = \sup_y\{\langle t\,x_1^* + (1-t)\,x_2^*, x \rangle - \varphi(x)\}$$

$$\geq t\,\sup_y\{\langle x_1^*, x \rangle - \varphi(x)\} + (1-t)\,\sup_y\{\langle x_2^*, x \rangle - \varphi(x)\}$$

$$= t\,\varphi^*(x_1^*) + (1-t)\,\varphi^*(x_2^*).$$

Thus, φ^* is convex. For $\epsilon > 0$ arbitrary let $y \in X$ be

$$\varphi^*(x^*) - \epsilon \geq \langle x^*, y \rangle - \varphi(y)$$

For $x_n \to x^*$

$$\varphi^*(x_n^*)\langle x_n^*, y \rangle - \varphi(y)$$

and

$$\liminf_{n\to\infty} \varphi^*(x_n^*) \geq \langle x^*, y \rangle - \varphi(y)$$

Since $\epsilon >$ arbitrary, φ^* is lower semi-continuous. $\square$

Definition (Bi-Conjugate Function) For any proper functional φ on X the bi-conjugate function of φ is $\varphi^{**} : X \to (-\infty, \infty]$ is defined by

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

Theorem C.3 For any $F : X \to (-\infty, \infty]$ F^{**} is convex and l.s.c. and $F^{**} \leq F$.

Proof: Note that

$$F^{**}(\bar{x}) = \sup_{x^* \in D(F^*)} \{\langle x^*, \bar{x} \rangle - F^*(x^*)\}$$

Since for $x^* \in D(F^*)$

$$\langle x^*, x \rangle - F(x^*) \leq F(x)$$

for all x, $F^{**}(\bar{x}) \leq F(\bar{x})$. $\square$

Theorem If F is a proper l.s.c convex function, then $F = F^{**}$. If X is reflexive, then $(F^*)^* = F^{**}$.

Proof: By Theorem if F is a proper l.s.c. convex function, then F is bounded below by an affine functional $\langle x^*, x \rangle - a$, i.e. $a \geq \langle x^*, x \rangle - F(x)$ for all $x \in X$. Define

$$a(x^*) = \inf\{a \in R : a \geq \langle x^*, x \rangle - F(x) \text{ for all } x \in X\}$$

Then for $a \geq a(x^*)$

$$\langle x^*, x \rangle - a \leq \langle x^*, x \rangle - a(x^*)$$

for all $x \in X$. By Theorem again

$$F(\bar{x}) = \sup_{x^* \in X^*} \{\langle x^*, \bar{x} \rangle - a(x^*)\}.$$

Since

$$a(x^*) = \sup_{x \in X} \{\langle x^*, x \rangle - F(x)\} = F^*(x)$$

we have

$$F(x) = \sup_{x^* \in X^*} \{\langle x^*, x \rangle - F^*(x^*)\}.\square$$

We have the duality property.

Theorem 2.2 Let F be a proper convex function on X.
(1) For $F : X \to (-\infty, \infty]$ $x^* \in \partial F(\bar{x})$ is if and only if

$$F(\bar{x}) + F^*(x^*) = \langle x^*, \bar{x} \rangle.$$

(2) Assume X is reflexive. If $x^* \in \partial F(\bar{x})$, then $\bar{x} \in \partial F^*(x^*)$. If F is convex and lower semi-continuous, then $x^* \in \partial F(\bar{x})$ if and only if $\bar{x} \in \partial F^*(x^*)$.

Proof: (1) Note that $x^* \in \partial F(\bar{x})$ is if and only if

$$\langle x^*, x\rangle - F(x) \leq \langle x^*, \bar{x}\rangle - F(\bar{x}) \tag{7.1}$$

for all $x \in X$. By the definition of F^* this implies $F^*(x^*) = \langle x^*, \bar{x}\rangle - F(\bar{x})$. Conversely, if $F(x^*) + F(\bar{x}) = \langle x^*, \bar{x}\rangle$ then (7.1) holds for all $x \in X$.
(2) Since $F^{**} \leq F$ by Theorem C.3, it follows from (1)

$$F^{**}(\bar{x}) \leq \langle x^*, \bar{x}\rangle - F^*(x^*)$$

But by the definition of F^{**}

$$F^{**}(\bar{x}) \geq \langle x^*, \bar{x}\rangle - F^*(x^*).$$

Hence

$$F^*(x^*) + F^{**}(\bar{x}) = \langle x^*, \bar{x}\rangle.$$

Since X is reflexive, $F^{**} = (F^*)^*$. If we apply (1) to F^* we have $\bar{x} \in \partial F^*(x^*)$. In addition if F is convex and l.s.c, it follows from Theorem C.1 that $F^{**} = F$. Thus if $x^* \in \partial F^*(\bar{x})$ then

$$F(\bar{x}) + F^*(x^*) = F^*(x^*) + F^{**}(x) = \langle x^*, \bar{x}\rangle$$

by applying (1) for F^*. Therefore $x^* \in \partial F(\bar{x})$ again by (1). □

Theorem(Rockafellar) Let X be real Banach space. If φ is lower-semicontinuous proper convex functional on X, then $\partial\varphi$ is a maximal monotone operator from X into X^*.

Proof: We prove the theorem when X is reflexive. By Apuland theorem we can assume that X and X^* are strictly convex. By Minty-Browder theorem $\partial\varphi$ it suffices to prove that $R(F + \partial\varphi) = X^*$. For $x_0^* \in X^*$ we must show that equation $x_0^* \in Fx + \partial\varphi(x)$ has at least a solution x_0 Define the proper convex functional on X by

$$f(x) = \frac{1}{2}\,|x|_X^2 + \varphi(x) - \langle x, x_0^*\rangle.$$

Since f is lower-semicontinuous and $f(x) \to \infty$ as $|x| \to \infty$ there exists $x_0 \in D(f)$ such that $f(x_0) \leq f(x)$ for all $x \in X$. Since F is monotone

$$\varphi(x) - \varphi(x_0) \geq \langle x - x_0, x_0^*\rangle - \langle x - x_0, F(x)\rangle$$

Setting $x_t = x_0 + t\,(u - x_0)$ and since φ is convex, we have

$$\varphi(u) - \varphi(x_0) \geq \frac{1}{t}\varphi(x_t) - \varphi(x_0) \geq \langle u - x_0, x_0^*\rangle - \langle u - x_0, F(x_t)\rangle.$$

Taking limit $t \to 0^+$, we obtain

$$\varphi(u) - \varphi(x_0) \geq \langle u - x_0, x_0^*\rangle - \langle u - x_0, F(x_0)\rangle,$$

which implies $x_0^* - F(x_0) \in \partial\varphi(x_0)$. □

We have the perturbation result.

<u>Theorem</u> Assume that X is a real Hilbert space and that A is a maximal monotone operator on X. Let φ be a proper, convex and lower semi-continuous functional on X satisfying $dom(A) \cap dom(\partial\varphi)$ is not empty and

$$\varphi((I+\lambda\,A)^{-1}x) \leq \varphi(x) + \lambda\,M, \qquad \text{for all } \lambda > 0,\ x \in D(\varphi),$$

where M is some non-negative constant. Then the operator $A+\partial\varphi$ is maximal monotone.

<u>Lemma</u> Let A and B be m-dissipative operators on X. Then for every $y \in X$ the equation

$$y \in -Ax - B_\lambda x \tag{7.2}$$

has a unique solution $x \in dom\,(A)$.

Proof Equation (9.17) is equivalent to $y = x_\lambda - w_\lambda - B_\lambda x_\lambda$ for some $w_\lambda \in A(x_\lambda)$. Thus,

$$x_\lambda - \frac{\lambda}{\lambda+1} w_\lambda = \frac{\lambda}{\lambda+1} y + \frac{1}{\lambda+1}(x_\lambda + \lambda\,B_\lambda x_\lambda)$$

$$= \frac{\lambda}{\lambda+1} y + \frac{1}{\lambda+1}(I-\lambda\,B)^{-1}.$$

Since A is m-dissipative, we conclude that (9.17) is equivalent to that x_λ is the fixed point of the operator

$$\mathcal{F}_\lambda x = (I - \frac{\lambda}{\lambda+1}\,A)^{-1}(\frac{\lambda}{\lambda+1}\,y + \frac{1}{\lambda+1}\,(I-\lambda\,B)^{-1}x).$$

By m-dissipativity of the operators A and B their resolvents are contractions on X and thus

$$|\mathcal{F}_\lambda x_1 - \mathcal{F}_\lambda x_2| \leq \frac{\lambda}{\lambda+1}\,|x_1 - x_2| \text{ for all } \lambda > 0,\ \ x_1,\ x_2 \in X.$$

Hence, $\mathcal{F}_\lambda$ has the unique fixed point x_λ and $x_\lambda \in dom\,(A)$ solves (9.17). □

Proof: From Lemma there exists x_λ for $y \in X$ such that

$$y \in x_\lambda - (-A)_\lambda x_\lambda + \partial\varphi(x_\lambda)$$

Moreover, one can show that $|x_\lambda|$ is bounded uniformly. Since

$$y - x_\lambda + (-A)_\lambda x_\lambda \in \partial\varphi(x_\lambda)$$

for $z \in X$

$$\varphi(z) - \varphi(x_\lambda) \geq (z - x_\lambda, y - x_\lambda + (-A)_\lambda x_\lambda)$$

Letting $\lambda(I+\lambda A)^{-1}x$, so that $z - x_\lambda = \lambda(-A)_\lambda x_\lambda$ and we obtain

$$(\lambda(-A)_\lambda x_\lambda, y - x_\lambda + (-A)_\lambda x_\lambda) \leq \varphi((I+\lambda\,A)^{-1}) - \varphi(x_\lambda) \leq \lambda\,M,$$

and thus

$$|(-A)_\lambda x_\lambda|^2 \leq |(-A)_\lambda x_\lambda||y - x_\lambda| + M.$$

Since $x_\lambda|$ is bounded and so that $|(-A)_\lambda x_\lambda|$.

7.2 Duality Theory

In this section we discuss the duality theory. We call

$$(P) \qquad \inf_{x\in X} \quad F(x)$$

is the primal problem, where X is a Banach space and $F : X \to (-\infty, \infty]$ is a proper l.s.c. convex function. We have the following result for the existence of minimizer.

Theorem 2 Let X be reflexive and F be a lower semicontinuous proper convex functional defined on X. Suppose F

$$\lim F(x) = \infty \quad \text{as } \ |x| \to \infty. \tag{7.3}$$

Then there exists an $\bar{x} \in X$ such that

$$F(\bar{x}) = \inf \{F(x); \;\; x \in X\}.$$

Proof: Let $\eta = \inf\{F(x); x \in X\}$ and let $\{x_n\}$ be a minimizing sequence such that $\lim F(x_n) = \eta$. The condition (7.3) implies that $\{x_n\}$ is bounded in X. Since X is reflexive there exists a subsequence that converges weakly to $\bar{x}$ in X and it follows from Lemma 1 that $F(\bar{x}) = \eta$. $\square$

We imbed (P) into a family of perturbed problem

$$(P_y) \qquad \inf_{x\in X} \quad \Phi(x, y)$$

where $y \in Y$, a Banach space is an imbedding variable and $\Phi : X \times Y \to (-\infty, \infty]$ is a proper l.s.c. convex function with $\Phi(x, 0) = F(x)$. Thus $(P_0) = (P)$. For example in terms of (12.1) we let

$$F(x) = f(x) + \varphi(\Lambda x) \tag{7.4}$$

and with $Y = H$

$$\Phi(x, y) = f(x) + \varphi(\Lambda x + y) \tag{7.5}$$

Definition The dual problem of (P) with respect to Φ is

$$(P^*) \qquad \sup_{y^*\in Y} (-\Phi^*(0, y^*)).$$

The value function of (P_y) is defined by

$$h(y) = \inf_{x\in X} \quad \Phi(x, y), \quad\; y \in Y.$$

We assume that $h(y) > -\infty$ for all $y \in Y$. Then we have

Theorem D.0 $\inf\,(P) \ge \sup\,(P^*)$.

Proof. For any $(x, y) \in X \times Y$ and $(x^*, y^*) \in X^* \times Y^*$

$$\Phi^*(x^*, y^*) \ge \langle x^*, x\rangle + \langle y^*, y\rangle - \Phi(z).$$

Thus

$$F(x) + \Phi^*(0, y^*) \ge \langle 0, x\rangle + \langle y^*, 0\rangle = 0$$

for all $x \in X$ and $y^* \in Y^*$. Therefore

$$\inf\,(P) = \inf\,F(x) \geq \sup_{y^* \in Y^*}\,(-\Phi^*(0, y^*)) = \sup\,(P^*).\square$$

In what follows we establish the equivalence of the primal problem (P) and the dual problem (P^*). We start with the properties of h.

Lemma D.1 h is convex.

Proof. The proof is by contradiction. Suppose there exist $y_1,\ y_2 \in Y$ and $\theta \in (0,1)$ such that

$$h(\theta\,y_1 + (1-\theta)\,y_2) > \theta\,h(y_1) + (1-\theta)\,h(y_2).$$

Then there exists c and $\epsilon > 0$ such that

$$h(\theta\,y_1 + (1-\theta)\,y_2) > c > c - \epsilon > \theta\,h(y_1) + (1-\theta)\,h(y_2).$$

Let $a_1 = h(y_1) + \frac{\theta}{\epsilon} > h(y_1)$ and

$$a_2 = \frac{c - \theta\,a_1}{1-\theta} = \frac{c - \epsilon - \theta\,h(y_)}{1-\theta} > h(y_2).$$

By definition of h there exist $x_1,\ x_2 \in X$ such that

$$h(y_1) \leq \Phi(x_,y_1) \leq a_1 \quad \text{and} \quad h(y_2) \leq \Phi(x_2, y_2) \leq a_2.$$

Thus

$$h(\theta\,y_1 + (1-\theta)\,y_2) \leq \Phi(\theta\,x_1 + (1-\theta)\,x_2, \theta\,y_1 + (1-\theta)\,y_2) \leq \theta\,\Phi(x_1, y_1) + (1-\theta)\,\Phi(x_2, y_2)$$

$$\leq \theta\,a_1 + (1-\theta)\,a_2 = c$$

which is a contradiction. Hence h is convex. $\square$

Lemma D.2 For all $y^* \in Y^*$, $h^*(y) = \Phi^*(0, y^*)$.

Proof.

$$h(y^*) = \sup_{y \in Y}\,(\langle y^*, y\rangle - h(y)) = \sup_{y \in Y}\,(\langle y^*, y\rangle - \inf_{x \in X}\,\Phi(x, y))$$

$$= \sup_{y \in Y}\sup_{x \in X}\,((\langle y^*, y\rangle - \Phi(x, y)) = \sup_{y \in Y}\sup_{x \in X}\,(\langle 0, x\rangle + \langle y^*, y\rangle - \Phi(x, y))$$

$$= \sup_{(x,y) \in X \times Y}(\langle (0, y^*), (x, y)\rangle - \Phi(x, y)) = \Phi^*(0, y^*).\square$$

Note that h is not necessarily l.s.c..

Theorem D.3 If h is l.s.c. at 0, then $\inf\,(P) = \sup\,(P^*)$.

Proof. Since F is proper, $h(0) = \inf_{x \in X}\,F(x) < \infty$. Since h is convex by Lemma D.1, it follows from Theorem C.4 that $h(0) = h^{**}(0)$. Thus

$$\sup\,(P^*) = \sup_{y^* \in Y^*}\,(-\Phi(0, y^*)) = \sup_{y^* \in Y^*}\,((\langle y^*, 0\rangle - h^*(y^*))$$

$$= h^{**}(0) = h(0) = \inf\,(P) square$$

Theorem D.4 If h is subdifferentiable at 0, then inf $(P) =$ sup (P^*) and $\partial h(0$ is the set of solutions of (P^*).

Proof. By Lemma D.2 $\bar{y}^*$ solves (P^*) if and only if

$$-h^*(\bar{y}^*) = -\Phi(0, \bar{y}^*) = \sup_{y^* \in Y^*} (-\Phi(0, y^*))$$

$$= \sup_{y^*} (\langle y^*, 0\rangle - h^*(y^*)) = h^{**}(0).$$

By Theorem 2.2

$$h^{**}(0) + h^{***}(\bar{y}^*) = \langle \bar{y}^*, 0\rangle = 0$$

if and only if $\bar{y}^* \in \partial h^{**}(0)$. Since $h^{***} = h^*$ by Theorem C.5, y^* solves (P^*) if and only if $y^* \in \partial h^{**}(y^*)$. Since $\partial h(0)$ is not empty, $\partial h(0) = \partial h^{**}(0)$. Therefore $\partial h(0)$ is the set of all solutions of (P^*) and (P^*) has at least one solution.

Let $y^* \in \partial h(0)$. Then

$$\langle y^*, x\rangle + h(0) \leq h(x)$$

for all $x \in X$. Let $\{x_n\}$ is a sequence in X such that $x_n \to 0$.

$$\liminf_{n\to\infty} h(x_n) \geq \lim \langle y^*, x_n\rangle + h(0) = h(0)$$

and h is l.s.c. at 0. By Theorem D.3 inf $(P) =$ sup (P^*). $\square$

Corollary D.6 If there exists an $\bar{x} \in X$ such that $\Phi(\bar{x}, \cdot)$ is finite and continuous at 0, h is continuous on an open neighborhood U of 0 and $h = h^{**}$. Moreover,

$$\inf\ (P) = \sup\ (P^*)$$

and $\partial h(0)$ is the set of solutions of (P^*).

Proof. First show that h is continuous. Clearly, $\Phi(\bar{X}, \cdot)$ is bounded above on an open neighborhood U of 0. Since for all $y \in Y$

$$h(y) \leq \Phi(\bar{x}, y),$$

$h$is bounded above on U. Since h is convex by Lemma D.1 and h is continuous by Theorem C.6. Hence $h = h^{**}$ by Theorem C. Now, h is subdifferential at 0 by Theorem C.10. The conclusion follows from Theorem D.4. $\square$

Example D.4 Consider the example (7.4)–(7.5), i.e.,

$$\Phi(x, y) = f(x) + \varphi(\Lambda x + y)$$

for (12.1). Let us calculate the conjugate of Φ.

$$\Phi^*(x^*, y^*) = \sup_{x\in X} \sup_{y\in Y} \{\langle x^*, x\rangle + \langle y^*, y\rangle - \Phi(x, y)\}$$

$$= \sup_{x\in X} \sup_{y\in Y} \{\langle x^*, x\rangle + \langle y^*, y\rangle - f(x) - \varphi(\Lambda x + y)\}$$

$$= \sup_{x\in X} \{\langle x^*, x\rangle - f(x) \sup_{y\in Y} [\langle y^*, y\rangle - \varphi(\Lambda x + y)]\}$$

where

$$\sup_{y\in Y} [\langle y^*, y\rangle - \varphi(\Lambda x + y)] = \sup_{y\in Y} [\langle y^*, \Lambda x + y\rangle - \varphi(\Lambda x + y) - \langle y^*, \Lambda x\rangle]$$

$$= \sup_{z\in Y} [\langle y^*, z\rangle - G(z)] - \langle y^*, \Lambda x\rangle = \varphi(y^*) - \langle y^*, \Lambda x\rangle.$$

Thus

$$\Phi^*(x^*,y^*) = \sup_{x\in X}\{\langle x^*,x\rangle - \langle y^*,\Lambda x\rangle - f(x) + \varphi(y^*)\}$$

$$= \sup_{x\in X}\{\langle x^* - \Lambda^* y^*, x\rangle - f(x) + \varphi(y^*)\} = f^*(x^* - \Lambda y^*) + \varphi^*(y^*).$$

Theorem D.7 For any $\bar{x} \in X$ and $\bar{y}^* \in Y^*$, the following statements are equivalent.

(1) $\bar{x}$ solves (P), $\bar{y}^*$ solves (P^*), and $\min\,(P) = \max\,(P^*)$.
(2) $\Phi(\bar{x}) + \Phi^*(0,\bar{y}^*) = 0$.
(3) $(0,\bar{y}^*) \in \partial\Phi(\bar{x},0)$.

Proof: Suppose (1) holds, then obviously (2) holds. Suppose (2) holds, then since

$$\Phi(\bar{x},0) = F(\bar{x}) \geq \inf\,(P) \geq \sup\,(P^*) \geq -\Phi(0,\bar{y}^*)$$

by Theorem D.2, thus

$$\Phi(\bar{x},0) = \min\,(P) = \sup\,(P^*) = -\Phi^*(0,\bar{y}^*)$$

Therfore (2) implies (1). Since $\langle (0,\bar{y}^*),(\bar{x},0)\rangle = 0$, (2) and (3) by Theorem C.7. $\square$

Any solution y^* of (P^*) is called a Lagrange multiplier associated with Φ. In fact for the above Example the optimality condition implies

$$\begin{aligned} 0 &= \Phi(\bar{x},0) + \Phi^*(0,\bar{y}^*) = F(\bar{x}) + F^*(-\Lambda^*\bar{y}^*) + \varphi(\Lambda\bar{x}) + \varphi(\bar{y}^*) \\ &= [F(\bar{x}) + F^*(-\Lambda^*\bar{y}^*) - \langle -\Lambda^*\bar{y}^*,\bar{x}\rangle] + [\varphi(\Lambda\bar{x}) + \varphi^*(\bar{y}^*) - \langle \bar{y}^*,\Lambda\bar{x}\rangle] \end{aligned} \tag{7.6}$$

Since each expression of (7.6) in square brackets is nonnegative, it follows that

$$F(\bar{x}) + F^*(-\Lambda^*\bar{y}^*) - \langle -\Lambda^*\bar{y}^*,\bar{x}\rangle = 0$$

$$\varphi(\Lambda\bar{x}) + \varphi^*(\bar{y}^*) - \langle \bar{y}^*,\Lambda\bar{x}\rangle = 0$$

By Theorem 2.2

$$\begin{aligned} &0 \in \partial F(\bar{x}) + \Lambda^*\bar{y}^* \\ &\bar{y}^* \in \partial\varphi(\Lambda\bar{x}). \end{aligned} \tag{7.7}$$

The function $L : X \times Y^* \to (-\infty,\infty]$ definded by

$$-L(u,y^*) = \sup_{y\in Y}\,\{\langle y^*,y\rangle - \Phi(x,y)\} \tag{7.8}$$

is called the Lagrangian. Note that

$$\Phi^*(x^*,y^*) = \sup_{x\in X,\; y\in Y}\,\{\langle x^*,x\rangle + \langle x^*,x\rangle - \Phi(x,y)\}$$

$$= \sup_{x\in X}\,\langle x^*,x\rangle + \sup_{y\in Y}\,\{\langle x^*,x\rangle - \Phi(x,y)\}\} = \sup_{x\in X}\,\langle x^*,x\rangle - L(x,y^*)$$

Thus

$$-\Phi^*(0,y^*) = \inf_{x\in X}\; L(x,y^*) \tag{7.9}$$

and thus the dual problem (P^*) is equivalently written as

$$\sup_{y^*\in Y^*}\ \inf_{x\in X}\ L(x,y^*).$$

Similarly, if Φ is a proper convex l.s.c. function, then the bi-conjugate of Φ in y, $\Phi_x^{**}=\Phi$ and

$$\Phi(x,y)=\Phi_x^{**}(y)=\sup\nolimits_{y^*\in Y^*}\ \{\langle y^*,y\rangle-\Phi_x^*(y^*)\}$$

$$=\sup_{y^*\in Y^*}\ \{\langle y^*,y\rangle+L(u,y^*)\}.$$

Hence

$$\Phi(u,0)=\sup_{y^*\in Y^*}\ L(x,y^*) \tag{7.10}$$

and the primal problem (P) is equivalently written as

$$\inf_{x\in X}\ \sup_{y^*\in Y}\ L(x,y^*).$$

By introducing the Lagrangian L, the Problems (P) and (P^*) are formulated as min-max problems, which arise in the game theory. By Theorem D.0

$$\sup_{y^*}\inf_{x}\ L(x,y^*)\le\inf_{x}\sup_{y^*}\ L(x,y^*).$$

Theorem D.8 (Saddle Point) Assume Φ is a proper convex l.s.c. function. Then the followings are equivalent.
(1) $(\bar{x},\bar{y}^*)\in X\times Y^*$ is a saddle point of L, i.e.,

$$L(\bar{u},y^*)\le L(\bar{u},\bar{y}^*)\le L(x,\bar{y}^*)\quad\text{for all } x\in X,\ y^*\in Y^*. \tag{7.11}$$

(2) $\bar{x}$ solves (P), $\bar{y}^*$ solves (P^*), and min $(P)=$ max (P^*).

Proof: Suppose (1) holds. From (7.9) and (7.11)

$$L(\bar{x},\bar{y}^*)=\inf_{x\in X}L(x,\bar{y}^*)=-\Phi^*(0,\bar{y}^*)$$

and form (7.10) and (7.11)

$$L(\bar{x},\bar{y}^*=\sup_{y^*\in Y^*}L(\bar{x},y^*)=\Phi(\bar{x},0)$$

Thus, $\Phi^*(\bar{u},0)+\Phi(0,\bar{y}^*)=0$ and (2) follows from Theorem D.7.

Conversely, if (2) holds, then from (7.9) and (7.10)

$$-\Phi^*(0,\bar{y}^*)=\inf\nolimits_{x\in X}\ L(x,\bar{y}^*)\le L(\bar{x},\bar{y}^*)$$

$$\Phi^*(\bar{x},0)=\sup\nolimits_{y^*\in Y^*}\ L(\bar{x}x,y)\ge L(\bar{x},\bar{y}^*)$$

By Theorem D.7 $-\Phi^*(0,\bar{y}^*=\Phi^*(\bar{x},0)$ and (7.11) holds. $\square$

Theorem D.8 implies that no duality gup between (P) and (P^*) is equivalent to the saddle point property of the pair $(\bar{x},\bar{y}^*)$.

For the Example

$$L(x, y^*) = f(x) + \langle y^*, \Lambda x\rangle - \varphi(y^*) \tag{7.12}$$

If $(\bar{x}, \bar{y}^*)$ is a saddle point, then from (7.11)

$$\begin{aligned} &0 \in \partial f(\bar{x}) + \Lambda^* \bar{y}^* \\ &0 \in \Lambda \bar{x} - \partial \varphi^*(\bar{x}). \end{aligned} \tag{7.13}$$

It follows from Theorem 2.2 that the second equation is equivalent to

$$\bar{y}^* \in \partial \varphi(\Lambda \bar{x})$$

and (7.13) is equivalent to (7.7).

7.3 Monotone Operators and Yosida-Morrey approximations

7.4 Lagrange multiplier Theory

In this action we introduce the generalized Lagrange multiplier theory. We establish the conditions for the existence of the Lagrange multiplier $\bar{\lambda}$ and derive the optimality condition (1.7)–(1.8). In Section 1.5 it is shown that the both Uzawa and augmented Lagrangian method are fixed point methods for the complementarity condition (**??**))and the convergence analysis is presented . In Section 1.6 we preset the concrete examples and demonstrate the applicability of the results in Sections 1.4–1.5.

8 Optimization Algorithms

In this section we discuss iterative methods for finding a solution to the optimization problem. First, consider the unconstrained minimization:

$$\min \quad J(x) \quad \text{over a Hilbert space } X.$$

Assume $J : X \to R^+$ is continuously differentiable. If $u^* \in X$ minimizes J, then the necessarily optimality is given by $J'(u^*) = 0$. The gradient method is given by

$$x_{n+1} = x_n - \alpha\, J'(x_n) \tag{8.1}$$

where $\alpha > 0$ is a stepsize chosen so that the decent $J(x_{n+1}) < J(x_n)$ holds. Let $g_n = J'(x_n)$ and since

$$J(x_n - \alpha\, g_n) - J(x_n) = -\alpha\, |J'(x_n)|^2 + o(\alpha), \tag{8.2}$$

there exits a $\alpha > 0$ satisfies the decent property. A stepsize $\alpha > 0$ can be determined by

$$\min_{\alpha>0} J(x_n - \alpha\, g_n) \tag{8.3}$$

or by a line search method (e.g, Amijo-rule)

Next, consider the constrained minimization

$$\min \quad J(x) \quad \text{over } x \in \mathcal{C}. \tag{8.4}$$

where $\mathcal{C}$ is a closed convex set in X. It follows from Theorem 7 that if $J : X \to R$ is C^1 and $x^* \in \mathcal{C}$ is a minimizer, then

$$(J'(x^*), x - x^*) \geq 0 \text{ for all } x \in \mathcal{C}. \tag{8.5}$$

Define the orthogonal projection operator $P_{\mathcal{C}}$ of X onto $\mathcal{C}$, i.e., $z^* = P_{\mathcal{C}}x$ minimizes

$$|z - x|^2 \text{ over } z \in \mathcal{C}.$$

From Theorem 7, $z^* \in \mathcal{C}$ satisfies

$$(z^* - x, z - z^*) \geq 0 \text{ for all } z \in \mathcal{C}. \tag{8.6}$$

Note that (8.5) is equivalent to

$$(x^* - (x^* - \alpha\, J'(x^*)), x - x^*) \geq 0 \text{ for all } x \in \mathcal{C} \text{ and } \alpha > 0.$$

Thus, it is follows from (8.6) that (8.5) is equivalent to

$$x^* = P_{\mathcal{C}}(x^* - \alpha\, J'(x^*)).$$

The projected gradient method for (8.4) is defined by

$$x_{n+1} = P_{\mathcal{C}}(x_n - \alpha_n\, J'(x_n)) \tag{8.7}$$

with a stepsize $\alpha_n > 0$.

Define the the Hessian $H = J''(x_0) \in \mathcal{L}(X, X)$ of J at x_0 by

$$J(x) - J(x_0) - (J'(x_0), x - x_0) = \frac{1}{2}\,(x - x_0, H(x - x_0)) + o(|x - x_0|^2).$$

Thus, we minimize over x

$$J(x) \sim J(x_0) + (J'(x_0), x - x_0) + \frac{1}{2}\,(x - x_0, H(x - x_0))$$

we obtain the damped-Newton method for equation given by

$$x_{n+1} = x_n - \alpha\, H(x_n)^{-1} J'(x_n), \tag{8.8}$$

where $\alpha > 0$ is a stepsize and $\alpha = 1$ corresponds to the Newton method. In the case of $\min \; J(x)$ over $x \in \mathcal{C}$ we obtain the projected Newton method

$$x_{n+1} = Proj_{\mathcal{C}}(x_n - \alpha\, H(x_n)^{-1} J'(x_n)). \tag{8.9}$$

Computing the Hessian can be expensive and numerical evaluation can also be less accurate. The finite rank scant update (BFGS and GFP) is used to substitute H (Quasi-Newton method). It is a variable merit method and such updates are preconditioner for the gradient method.

Consider the (regularized) nonlinear least square problem

$$\min \quad \frac{1}{2}|F(x)|_Y^2 + \frac{\alpha}{2}(Px, x)_X \tag{8.10}$$

where $F : X \to Y$ is C^1 and P is a nonnegative self-adjoint operator on a Hilbert space X. The Gauss-Newton method is an iterative method of the form

$$x_{n+1} = \text{argmin}\{\frac{1}{2}|F'(x_n)(x - x_n) + F(x_n)|_Y^2 + \frac{\alpha}{2}(Px, x)_X\},$$

i.e.,

$$x_{n+1} = x_n - (F'(x_n)^* F(x_n) + \alpha\, P)^{-1}(F'(x_n)^* F(x_n)). \tag{8.11}$$

8.1 Constrained minimization and Lagrange multiplier method

Consider the constrained minimization

$$\min \quad J(x, u) = F(x) + H(u) \quad \text{subject to } E(x, u) = 0 \text{ and } u \in \mathcal{C}. \tag{8.12}$$

We use the implicit function theory for developing algorithms for (8.12).

Implicit Function Theory Let $E : X \times U \to X$ is C^1. Suppose a pair $(\bar{x}, \bar{u})$ satisfies $E(x, u) = 0$ and $E_x(\bar{x}, \bar{u})$ is bounded invertible. Then the exists a $\delta > 0$ such that for $|u - \bar{u}| < \delta$ equation $E(x, u) = 0$ has a locally defined unique solution $x = \Phi(u)$. Moreover, $\Phi : U \to X$ is continuously differentiable at $\bar{u}$ and $\dot{x} = \Phi'(\bar{x})(d)$ satisfies

$$E_x(\bar{x}, \bar{u})\dot{x} + E_u(\bar{x}, \bar{u})d = 0.$$

Theorem (Lagrange Calculus) Assume that $E(x, u) = 0$ and $E_x(x, u)$ is bounded invertible. Then, by the implicit function theory $x = \Phi(u)$ and for $J(u) = J(\Phi(u), u)$

$$(J'(u), d) = (H'(u), d) + (E_u(x, u)^*\lambda, d)$$

where $\lambda \in Y$ satisfies

$$E_x(x, u)^*\lambda + F'(x) = 0.$$

Proof: From the implicit function theory

$$(J', d) = (F'(x), \dot{x}) + (H'(u), d),$$

where

$$E_x(x, u)\dot{x} + E_u(x, u) = 0.$$

Thus, the claim follows from

$$(F'(x), \dot{x}) = -(E_x^*\lambda, \dot{x}) = -(\lambda, E_x\dot{x}) = (\lambda, E_u d). \square$$

Hence we obtain the gradient method for equality constraint problem (8.12);

Gradient method

1. Solve for x_n

$$E(x_n, u_n) = 0$$

2. Solve for λ_n
$$E_x(x_n,u_n)^*\lambda_n + F'(x_n) = 0$$

3. Compute the gradient by
$$g_n = H'(u_n) + E_u(x_n,u_n)^*\lambda_n$$

4. Gradient Step: Update
$$u_{n+1} = P_{\mathcal{C}}(u_n - \alpha\, g_n)$$
with a line search for $\alpha > 0$.

8.2 Conjugate gradient method

In this section we develop the conjugate gradient and residual method for the saddle point problem.

Consider the quadratic programing
$$J(x) = \frac{1}{2}(Ax,x)_X - (b,x)_X$$
where A is positive self-adjioint operator on a Hilbert space X. Then, the negative gradient of J is given by
$$r = -J'(x) = b - Ax.$$
The gradient method is given by
$$x_{k+1} = x_k + \alpha_k\, p_k, \quad p_k = r_k,$$
where the step size α_k is selected by
$$\min \quad J(x_k + \alpha\, r_k) \text{ over } \alpha > 0.$$
That is, the Cauchy step is given by
$$0 = (J'(x_{k+1}), p_k) = \alpha\,(Ap_k,p_k) - (p_k,r_k) \Rightarrow \alpha_k = \frac{(p_k,r_k)}{(Ap_k,p_k)}. \tag{8.13}$$
It is known that the gradient method with Cauchy step searches the solution on a subspace spanned by the major and minor axes of A dominantly, and exhibits a zig-zag pattern of its iterates. If the condition number
$$\rho = \frac{\max\,\sigma(A)}{\min\,\sigma(A)}$$
is large, then it is very slow convergent. In order to improve the convergence we introduce the conjugacy of the search directions p_k and it can be achieved by the conjugate gradient method. For example, a new search direction p_k is extracted for the residual $r_k = b - Ax_k$ by
$$p_k = r_k - \beta\, p_{k-1}$$
and the orthogonality
$$0 = (Ap_k, p_{k-1}) = (r_k, Ap_{k-1}) - \beta(Ap_{k-1}, p_{k-1})$$

implies

$$\beta_k = \frac{(r_k, Ap_{k-1})}{(Ap_{k-1}, p_{k-1})}. \tag{8.14}$$

Let P is the pre-conditioner, self-adjiont operator so that the condition number of $P^{-\frac{1}{2}}AP^{-\frac{1}{2}}$ is much smaller that the one for A. If we apply (8.13)(8.14) with the pre-conditioner P, we have the pre-conditioned conjugate gradient method:

- Initialize x_0 and set

$$r_0 = b - Ax_0, \;\; z_0 = P^{-1}r_0, \;\; p_0 = z_0$$

- Iterate on k until (r_k, z_k) is sufficiently small

$$\alpha_k = \frac{(r_k, z_k)}{(p_k, Ap_k)}, \quad x_{k+1} = x_k + \alpha_k\, p_k, \quad r_{k+1} = r_k - \alpha_k\, Ap_k.$$

$$z_{k+1} = P^{-1}r_{k+1}, \;\; \beta_k = \frac{(z_{k+1}, r_{k+1})}{(z_k, r_k)}, \quad p_{k+1} := z_{k+1} + \beta_k\, p_k$$

Here, z_k is the pre-conditioned residual and $(r_k, z_k) = (r_k, P^{-1}r_k)$. The above formulation is equivalent to applying the conjugate gradient method without preconditioned system

$$P^{-\frac{1}{2}}AP^{-\frac{1}{2}}\hat{x} = P^{-\frac{1}{2}}b,$$

where $\hat{x} = P^{\frac{1}{2}}x$. We have the conjugate gradient theory.

Theorem (Conjugate gradient method)

(1) We have conjugate property

$$(r_k, z_j) = 0, \;\; (p_k, Ap_j) = 0 \text{ for } 0 \le j < k.$$

(2) x_{m+1} minimizes $J(x)$ over $x \in x_0 + span\{p_0, \cdots, p_m\}$.

Proof: For $k = 1$

$$(r_1, z_0) = (r_0, z_0) - \alpha_0\,(Az_0, z_0) = 0$$

and since $(r_1, z_1) = (r_0, z_1) - \alpha_0\,(r_1, Az_0)$

$$(p_1, Ap_0) = (z_1, Az_0) + \beta_0(z_0, Az_0) = -\frac{1}{\alpha_0}(r_1, z_1) + \beta_0(z_0, Az_0) = 0$$

Note that $span\{p_0, \cdots, p_k\} = span\{z_0, \cdots, z_k\}$. In induction in k, for $j < k$

$$(r_{k+1}, z_j) = (r_k, z_j) - \alpha_k(Ap_k, p_j) = 0$$

and

$$((r_{k+1}, z_k) = (r_k, z_k) - \alpha_k(Ap_k, p_k) = 0$$

Also, for $j < k$

$$(p_{k+1}, Ap_j) = (z_{k+1}, Ap_j) + \beta_k(p_k, Ap_j) = -\frac{1}{\alpha_j}(z_{k+1}, r_{j+1} - r_j) = 0$$

and

$$(p_{k+1}, Ap_k) = (z_{k+1}, Ap_k) + \beta_k(p_k, Ap_k) = -\frac{1}{\alpha_k}(r_{k+1}, z_k) + \beta_k(p_k, Ap_k) = 0.$$

Hence (1) holds. For (2) we note that

$$(J'(x_{m+1}), p_j) = (r_{m+1}, p_j) = 0 \text{ for } j \leq m$$

and $x_{m+1} = x_0 + span\{p_0, \cdots, p_m\}$. Thus, (2) holds. □

Remark Note that

$$r_k = r_0 - \sum_{j=0}^{k-1} \alpha_j P^{-1} A p_j$$

Thus,

$$|r_k|^2 = |r_0|^2 - (2 \sum^{k-1} \alpha_j P^{-1} A p_j, r_0) + |\alpha_j|.$$

Consider the constrained quadratic programing

$$\min \ \frac{1}{2}(Qy, y)_X - (a, y)_X \quad \text{subject to } Ey = b \text{ in } Y.$$

Then the necessary optimality condition is written for $x = (y, \lambda) \in X \times Y$:

$$A(y, \lambda) = (a, b) \text{ where } A = \begin{pmatrix} Q & E^* \\ & \\ E & 0 \end{pmatrix},$$

where A is self-adjoint but is indefinite.

In general we assume A is self-adjoint and invertible. Consider the least square problem

$$\frac{1}{2}|Ax - b|_X^2 = \frac{1}{2}(AAx, x)_X - (Ab, x)_X + \frac{1}{2}|b|^2.$$

We consider the conjugate directions $\{p_k\}$ satisfying $(Ap_k, Ap_j) = 0, \ j < k$. The preconditioned conjugate residual method may be derived in the same way as done for the conjugate gradient method:

- Initialize x_0 and set

$$r_0 = P^{-1}(b - Ax_0), \quad p_0 = r_0, \quad (Ap)_0 = Ap_0$$

- Iterate with k until (r_k, r_k) is sufficiently small

$$\alpha_k = \frac{(r_k, Ar_k)}{((Ap)_k, P^{-1}(Ap)_k)}, \quad x_{k+1} = x_k + \alpha_k \, p_k, \quad r_{k+1} = r_k - \alpha_k \, P^{-1}(Ap)_k.$$

$$\beta_k = \frac{(r_{k+1}, Ar_{k+1})}{(r_k, Ar_k)}, \quad p_{k+1} = r_{k+1} + \beta_k \, p_k, \quad (Ap)_{k+1} = Ar_{k+1} + \beta_k \, (Ap)_k$$

Theorem (Conjugate residual method) Assuming $\alpha_k \neq 0$,
(1) We have conjugate property

$$(r_j, Ar_k) = 0, \ \ (Ap_k, P^{-1}Ap_j) = 0 \text{ for } 0 \leq j < k.$$

(2) x_{m+1} minimizes $|Ax - b|^2$ over $x \in x_0 + span\{p_0, \cdots, p_m\}$.

Proof: For $k = 1$

$$(Ar_0, r_1) = (Ar_0, r_0) - \alpha_0\,(Ap_0, P^{-1}Ap_0) = 0$$

and since $(Ap_0, P^{-1}Ap_1) = (P^{-1}Ap_0, Ar_1) + \beta_0\,(Ap_0, P^{-1}Ap_0)$ and $(P^{-1}Ap_0, Ar_1) = \frac{1}{\alpha_0}(r_0 - r_1, Ar_1) = -\frac{1}{\alpha_0}(r_1, Ar_1)$, we have

$$(Ap_0, P^{-1}Ap_1) = -\frac{1}{\alpha_0}(r_1, Ar_1) + \beta_0(Ap_0, P^{-1}Ap_0) = 0.$$

Note that $span\{p_0, \cdots, p_k\} = span\{r_0, \cdots, r_k\}$ and $r_k = r_0 + P^{-1}A\,span\{p_0, \cdots, p_{k-1}\}$. In induction in k, for $j < k$

$$(r_{k+1}, Ar_j) = (r_k, Ar_j) - \alpha_k(P^{-1}Ap_k, Ar_j) = 0$$

and

$$(Ar_{k+1}, r_k) = (Ar_k, r_k) - \alpha_k(Ap_k, P^{-1}Ap_k) = 0$$

Also, for $j < k$

$$(Ap_{k+1}, P^{-1}Ap_j) = (Ar_{k+1}, P^{-1}Ap_j) + \beta_k(Ap_k, P^{-1}Ap_j) = -\frac{1}{\alpha_j}(Ar_{k+1}, r_{j+1} - r_j) = 0$$

and

$$(Ap_{k+1}, P^{-1}Ap_k) = (Ar_{k+1}, P^{-1}Ap_k) + \beta_k(Ap_k, P^{-1}Ap_k) = -\frac{1}{\alpha_k}(r_{k+1}, Ar_{k+1}) + \beta_k(p_k, Ap_k) = 0.$$

Hence (1) holds. For (2) we note that

$$(Ax_{m+1} - b, P^{-1}Ap_j) = -(r_{m+1}, P^{-1}Ap_j) = 0 \text{ for } j \le m$$

and $x_{m+1} = x_0 + span\{p_0, \cdots, p_m\}$. Thus, (2) holds. $\square$

direct inversion in the iterative subspace (DIIS)

8.3 Nonlinear Conjugate Residual method

One can extend the conjugate residual method for the nonlinear equation. Consider the equality constrained optimization

$$\min J(y) \text{ subject to } E(y) = 0.$$

The necessary optimality system for $x = (y, \lambda)$ is

$$F(y, \lambda) = \begin{pmatrix} J'(y) + E'(y)^*\lambda \\ \\ E(y) \end{pmatrix} = 0.$$

Multiplication of vector by the Jacobian F' can be approximated by the difference of two residual vectors by

$$\frac{F(x_k + t\,y) - F(x_k)}{t} \sim Ay \text{ with } t\,|y| \sim 1.$$

Thus, we have the nonlinear version of the conjugate residual method:

Nonlinear Conjugate Residual method

- Calculate s_k for approximating Ar_k by

$$s_k = \frac{F(x_k + t\, r_k) - F(x_k)}{t}, \quad t = \frac{1}{|r_k|}.$$

- Update the direction by

$$p_k = P^{-1} r_k - \beta_k\, p_{k-1}, \quad \beta_k = \frac{(s_k, r_k)}{(s_{k-1}, r_{k-1})}.$$

- Calculate

$$q_k = s_k - \beta_k\, q_{k-1}.$$

- Update the solution by

$$x_{k+1} = x_k + \alpha_k\, p_k, \quad \alpha_k = \frac{(s_k, r_k)}{(q_k, P^{-1} q_k)}.$$

- Calculate the residual

$$r_{k+1} = F(x_{k+1}).$$

9 Newton method and SQP

Next, we develop the Newton method for $J'(u) = 0$. Define the Lagrange functional

$$L(x, u, \lambda) = F(x) + H(u) + (\lambda, E(x, u)).$$

Note that

$$J'(u) = L_u(x(u), u, \lambda(u))$$

where $(x(u), \lambda(u))$ satisfies

$$E(x(u), u) = 0, \quad E'(x(u))^* \lambda(u) + F'(x(u)) = 0$$

where we assumed $E_x(x, u)$ is bounded invertible. By the implicit function theory

$$J''(u) = L_{ux}(x(u), u, \lambda(u))\dot{x} + E_u(x(u), u, \lambda(u))^* \dot{\lambda} + L_{uu}(x(u), u, \lambda(u)),$$

where $(\dot{x}, \dot{\lambda})$ satisfies

$$\begin{pmatrix} L_{xx}(x, u, \lambda) & E_x(x, u)^* \\ \\ E_x(x, u) & 0 \end{pmatrix} \begin{pmatrix} \dot{x} \\ \\ \dot{\lambda} \end{pmatrix} + \begin{pmatrix} L_{xu}(x, u, \lambda) \\ \\ E_u(x, u) \end{pmatrix} = 0.$$

Thus,

$$\dot{x} = -(E_x(x, u))^{-1} E_u(x, u), \quad \dot{\lambda} = -(E_x(x, u)^*)^{-1} (L_{xx}(x, u, \lambda)\dot{x} + L_{xu}(x, u, \lambda))$$

and

$$J''(u) = L_{ux}\dot{x} - E_u(x, u)^* (E_x(x, u)^*)^{-1} (L_{ux}\dot{x} + L_{xu}) + L_{uu}(x, u, \lambda).$$

Theorem (Lagrange Calculus II) The Newton step

$$J''(u)(u^+ - u) + J'(u) = 0$$

is equivalently written as

$$\left(\begin{array}{ccc} L_{xx}(x,u,\lambda) & L_{xu}(x,u,\lambda) & E_x(x,u)^* \\ \\ L_{ux}(x,u,\lambda) & L_{uu}(x,u,\lambda) & E_u(x,u)^* \\ \\ E_x(x,u) & E_u(x,u,\lambda) & 0 \end{array}\right)\left(\begin{array}{c} \Delta x \\ \\ \Delta u \\ \\ \Delta\lambda \end{array}\right)+\left(\begin{array}{c} 0 \\ \\ E_u(x,u)^*\lambda+H'(u) \\ \\ 0 \end{array}\right)=0$$

where $\Delta x=x^+-x$, $\Delta u=u^+-u$ and $\Delta\lambda=\lambda^+-\lambda$.

Proof: Define the linear operator T by

$$T(x,u)=\left(\begin{array}{c} -E_x(x,u)^{-1}E_u(x,u) \\ \\ I \end{array}\right).$$

Then, the Newton method is equivalently written as

$$T(x,u)^*L''(x,u,\lambda)\left(\begin{array}{c} \Delta x \\ \Delta u \end{array}\right)+T(x,u)^*\left(\begin{array}{c} 0 \\ E_u^*\lambda+H'(u) \end{array}\right)=0.$$

Since $E'(x,u)$ is surjective, it follows from the closed range theorem

$$N(T(x,u)^*)=R(T(x,u))^\perp=N(E'(x,u)^*)^\perp=R(E'(x,u)).$$

Hence, we obtain the claimed update. $\Box$

Thus, we have the Newton method for the equality constraint problem (8.12).

<u>Newton method</u>

1. Given u_0, initialize (x_0,λ_0) by

$$E(x_0,u_0)=0,\quad E_x(x_0,u_0)^*\lambda_0+F'(x_0)=0$$

2. Newton Step: Solve for $(\Delta x,\Delta u,\Delta\lambda)$

$$\left(\begin{array}{cc} L''(x_n,u_n,\lambda_n) & E'(x_n,u_n)^* \\ \\ E'(x_n,u_n) & 0 \end{array}\right)\left(\begin{array}{c} \Delta x \\ \Delta u \\ \Delta\lambda \end{array}\right)+\left(\begin{array}{c} 0 \\ E_u(x_n,u_n)^*\lambda_n+H'(u_n) \\ 0 \end{array}\right)=0$$

and update $u_{n+1}=Proj_{\mathcal{C}}(u_n+\alpha\,\Delta u)$ with a stepsize $\alpha>0$.

3. Feasible Step: Solve for (x_{n+1},λ_{n+1})

$$E(x_{n+1},u_{n+1})=0,\quad E_x(x_{n+1},u_{n+1})^*\lambda_{n+1}+F'(x_{n+1})=0.$$

4. Stop or set $n=n+1$ and return to step 2.

where

$$L''(x_n,u_n,\lambda_n)=\left(\begin{array}{cc} F''(x_n)+(\lambda_n,E_{xx}(x_n,u_n)) & (\lambda_n,E_{xu}(x_n,u_n)) \\ \\ (\lambda_n,E_{ux}(x_n,u_n)) & H''(u_n)+(\lambda_n,E_{uu}(x_n,u_n)) \end{array}\right).$$

and

$$E'(x_n,u_n) = (E_x(x_n,u_n), E_u(x_n,u_n)).$$

Consider the general equality constraint minimization;

$$\min \quad J(y) \quad \text{subject to} \quad E(y) = 0$$

The necessary optimality condition is given by

$$J'(y) + E'(y)^*\lambda = 0, \quad E(y) = 0, \tag{9.1}$$

where we assume $E'(y)$ is surjective. Problem (8.12) is a specific case of this, i.e.,

$$y = (x,u), \quad J(y) = F(x) + H(u), \quad E = E(x,u).$$

The Newton method applied to the necessary optimality (9.1) for (y,λ) is as follows.

One can apply the Newton method for the necessary optimality system

$$(L_x(x,u,\lambda), L_u(x,u,\lambda, E(x,u)) = 0$$

and results in the SQP (sequential quadratic programing):

Newton method applied to Necessary optimality system (SQP)

1. Newton direction: Solve for $(\Delta y, \Delta\lambda)$

$$\begin{pmatrix} L''(y_n,\lambda_n) & E'(y_n)^* \\ \\ E'(y_n) & 0 \end{pmatrix} \begin{pmatrix} \Delta y \\ \\ \Delta\lambda \end{pmatrix} + \begin{pmatrix} J'(y_n) + E'(y_n)^*\lambda_n \\ \\ E(y_n) \end{pmatrix} = 0. \tag{9.2}$$

2. Update $y_{n+1} = y_n + \alpha\,\Delta y$ and $\lambda_{n+1} = \lambda_m + \alpha\,\Delta\lambda$.
3. Stop or set $n = n+1$ and return to step 1.

where

$$L''(y,\lambda) = J''(y) + E''(y)^*\lambda.$$

The both Newton methods solve the same linear system (9.2) with the different right hand side. But, the first Newton method we iterate on the variable u only but each step uses the feasible step $E(y_n) = E(x_n,u_n) = 0$ and $L_x(x_n,u_n,\lambda_n) = 0$ for (x_n,λ_n), which are the steps for the gradient method. The, it computes the Newton direction for $J'(u)$ and updates u. In this sense the second Newton's method is a relaxed version of the first one.

9.1 Sequential programing

In this section we discuss the sequential programming. It is an intermediate method between the gradient method and the Newton method. Consider the constrained optimization

$$F(y) \text{ subject to } E(y) = 0, \quad y \in \mathcal{C}.$$

We linearize the equality constraint and consider a sequence of the constrained optimization;

$$\min \quad F(y) \text{ subject to } E'(y_n)(y-y_n) + E(y_n) = 0, \quad y \in \mathcal{C}. \tag{9.3}$$

The necessary optimality (if $E^(y_n)$ is surjective) is given by

$$\begin{cases} (F'(y)+E'(y_n)^*\lambda, \tilde{y}-y) \geq 0 \text{ for all } \tilde{y} \in \mathcal{C} \\ \\ E'(y_n)(y-y_n)+E(y_n)=0. \end{cases} \tag{9.4}$$

Note that the necessary condition for (y^*, λ^*) is written as

$$\begin{cases} (F'(y^*)+E'(y_n)^*\lambda^*-(E'(y_n)^*-E'(y^*)^*)\lambda^*, \tilde{y}-y^*) \geq 0 \text{ for all } \tilde{y} \in \mathcal{C} \\ \\ E'(y_n)(y^*-y_n)+E(y_n)=E'(y_n)(y^*-y_n)+E(y_n)-E(y^*). \end{cases} \tag{9.5}$$

Suppose we have the Lipschitz continuity of solutions to (9.4) and (9.5) with respect to the perturbation; $\Delta_1=(E'(y_n)^*-E'(y^*)^*)\lambda^*$ and $\Delta_2=E'(y_n)(y^*-y_n)+E(y_n)-E(y^*)$, we have

$$|y-y^*| \leq c_1\,|(E'(y_n)^*-E'(y^*)^*)\lambda^*|+c_2\,|E'(y_n)(y^*-y_n)+E(y_n)-E(y^*)|.$$

Thus, for the (damped) update:

$$y_{n+1}=(1-\alpha)y_n+\alpha\, y, \quad \alpha \in (0,1), \tag{9.6}$$

we have the estimate

$$|y_{n+1}-y^*| \leq (1-\alpha+\alpha\gamma)|y_n-y^*|+c\alpha\,|y_n-y^*|^2, \tag{9.7}$$

for some constants γ and $c>0$. In fact, for the case $\mathcal{C}=X$ let $z_n=(y_n,\lambda_n)$, $z^*=(y^*,\lambda^*)$ and define G_n by

$$G_n=\begin{pmatrix} F''(y_n) & E'(y_n)^* \\ \\ E'(y_n) & 0 \end{pmatrix}.$$

For the update:

$$z_{n+1}=(1-\alpha)z_n+\alpha\, z, \quad \alpha \in (0,1]$$

we have

$$z_{n+1}-z^*=(1-\alpha)\,(z_n-z^*)+\alpha\, G_n^{-1}(\begin{pmatrix} (E'(y_n)^*-E'(y^*)^*)\lambda^* \\ 0 \end{pmatrix}+\delta_n)$$

with

$$\delta_n=\begin{pmatrix} F'(y^*)-F'(y_{n+1})-F''(y_n)(y^*-y_{n+1}) \\ E(y^*)-E'(y_n)(y^*-y_n)-E(y_n). \end{pmatrix}.$$

Note that if $F''(y_n)=A$, then

$$G_n^{-1}\begin{pmatrix} \Delta_1 \\ 0 \end{pmatrix}=\begin{pmatrix} A^{-1}\Delta_1-E'(y_n)^*e \\ e \end{pmatrix}, \quad e=(E'(y_n)A^{-1}E'(y_n)^*)^{-1}E'(y_n)A^{-1}\Delta_1$$

Thus, we have

$$|G_n^{-1}\begin{pmatrix} (E'(y_n)^*-E'(y^*)^*)\lambda^* \\ 0 \end{pmatrix}| \leq \gamma\,|y^n-y^*|$$

Since

$$|G_n^{-1}\delta_n| \leq c\,|y_n - y^*|^2,$$

we have

$$|z_{n+1} - z^*| \leq (1 - \alpha + \alpha\gamma)\,|z_n - z^*| + \alpha c\,|z_n - z^*|^2$$

which imply (9.7). When either the quadratic variation of E is dominated by the linearization E' of E, i.e.,

$$|E'(y^*)^\dagger(E'(y^n) - E'(y^*))| \leq \beta\,|y^n - y|$$

with small $\beta > 0$ or the multiplier $|\lambda^*|$ is small, $\gamma > 0$ is sufficiently small. The estimate (9.7) implies that if $|y - y^*|^2$ is dominating we use a small $\alpha > 0$ to the damp out and if $|y - y^*|$ is dominating we use $\alpha \to 1$.

9.1.1 Second order version

The second order sequential programming is given by

$$\begin{aligned} \min \quad & F(y) + \langle \lambda_n, E(y) - (E'(y_n)(y - y_n) + E(y_n))\rangle \\ & \text{subject to } E'(y_n)(y - y_n) + E(y_n) = 0 \text{ and } y \in \mathcal{C}. \end{aligned} \tag{9.8}$$

The necessary optimality condition for (9.8) is given by

$$\begin{cases} (F'(y) + (E'(y)^* - E'(y_n)^*)\lambda_n + E'(y_n)^*\lambda, \tilde{y} - y) \geq 0 \text{ for all } \tilde{y} \in \mathcal{C} \\ \\ E'(y_n)(y - y_n) + E(y_n) = 0. \end{cases} \tag{9.9}$$

It follows from (9.5) and (9.9) that

$$|y - y^*| + |\lambda - \lambda^*| \leq c\,(|y_n - y^*|^2 + |\lambda_n - \lambda^*|^2).$$

In summary we have

Sequential Programming I

1. Given $y_n \in \mathcal{C}$, we solve for

$$\min \quad F(y) \text{ subject to } E'(y_n)(y - y_n) + E(y_n) = 0, \quad \text{over } y \in \mathcal{C}.$$

2. Update $y_{n+1} = (1 - \alpha)\,y_n + \alpha\,y,\ \alpha \in (0, 1)$. Iterate until convergence.

Sequential Programming II

1. Given $y_n \in \mathcal{C},\ \lambda_n$ and we solve for $y \in \mathcal{C}$:

$$\min \quad F(y) + (\lambda_n, E(y) - (E'(y_n)(y - y_n) + E(y_n)) \text{ subject to } E'(y_n)(y - y_n) + E(y_n) = 0, \quad y \in \mathcal{C}.$$

2. Update $(y_{n+1}, \lambda_{n+1}) = (y, \lambda)$. Iterate until convergence.

9.1.2 Examples

Consider the constrained optimization of the form

$$\min \quad F(x)+H(u) \quad \text{subject to } E(x,u)=0 \text{ and } u\in\mathcal{C}. \tag{9.10}$$

Suppose $H(u)$ is nonsmooth. We linearize the equality constraint and consider a sequence of the constrained optimization;

$$\begin{aligned}&\min \quad F(x)+H(u)+\langle\lambda_n, E(x,u)-(E'(x_n,u_n)(x-x_n,u-u_n)+E(x_n,u_n)\rangle\\ &\qquad \text{subject to } E'(x_n,u_n)(x-x_n,u-u_n)+E(x_n,u_n)=0 \text{ and } u\in\mathcal{C}.\end{aligned} \tag{9.11}$$

Note that if (x^*,u^*) is an optimizer of (9.10), then

$$\begin{aligned}&F'(x^*)+(E'(x^*,u^*)-E'(x_n,u_n))\lambda_n+E'(x_n,u_n)\lambda^*=(E'(x_n,u_n)-E'(x^*,u^*))(\lambda^*-\lambda_n)=\Delta_1\\ &E'(x_n,u_n)(x^*-x_n,u^*-u_n)+E(x_n,u_n)=\Delta_2, \quad |\Delta_2|\sim M\,(|x_n-x^*|^2+|u_n-u^*|^2).\end{aligned} \tag{9.12}$$

Thus, (x^*,u^*) is the solution to the perturbed problem of (9.11):

$$\begin{aligned}&\min \quad F(x)+H(u)+\langle\lambda_n, E(x,u)-(E'(x_n,u_n)(x-x_n,u-u_n)+E(x_n,u_n)\rangle\\ &\qquad \text{subject to } E'(x_n,u_n)(x-x_n,u-u_n)+E(x_n,u_n)=\Delta \text{ and } u\in\mathcal{C}.\end{aligned} \tag{9.13}$$

Let (x_{n+1},u_{n+1}) be the solution of (9.11). One can assume the Lipschitz continuity of solutions to (9.13):

$$|x_{n+1}-x^*|+|u_{n+1}-u^*|+|\lambda_n-\lambda^*|\le C\,|\Delta| \tag{9.14}$$

From (9.12) and assumption (9.14) we have the quadratic convergence of the sequential programing method (9.17):

$$|x_{n+1}-x^*|+|u_{n+1}-u^*|+|\lambda-\lambda^*|\le M\,(|x_n-x^*|^2+|u_n-u^*|^2+|\lambda_n-\lambda^*|). \tag{9.15}$$

Assuming H is convex, the necessary optimality given by

$$\begin{cases} F'(x)+(E_x(x,u)^*-E_x(x_n,u_n)^*)\lambda_n+E_x(x_n,u_n)^*\lambda=0\\[1ex] H(v)-H(u)+((E_u(x,u)^*-E_u(x_n,u_n)^*)\lambda_n+E_u(x,u)^*\lambda, v-u)\ge 0 \text{ for all } v\in\mathcal{C}\\[1ex] E_x(x_n,u_n)(x-x_n)+E_u(x_n,u_n)(u-u_n)+E(x_n,u_n)=0.\end{cases} \tag{9.16}$$

As in the Newton method we use

$$\begin{aligned}&(E_x(x,u)^*-E_x(x_n,u_n)^*)\lambda_n\sim E''(x_n,u_n)(x-x_n,u-u_n)\lambda_n\\ &F'(x_{n+1})\sim F''(x_n)(x_{n+1}-x_n)+F'(x_n)\end{aligned}$$

and obtain

$$\begin{aligned}&F''(x_n)(x-x_n)+F'(x_n)+\lambda_n(E_{xx}(x_n,u_n)(x-x_n)+E_{xu}(x_n,u_n)(u-u_n))+E_x(x_n,u_n)^*\lambda=0\\ &E_x(x_n,u_n)x+E_u(x_n,u_n)u=E_x(x_n,u_n)x_n+E_u(x_n,u_n)u_n-E(x_n,u_n),\end{aligned}$$

which is a linear system of equations for (x,λ). Thus, one has $x = x(u)$, $\lambda = \lambda(u)$, a continuous affine function in u;

$$\begin{pmatrix} x(u) \\ \\ \lambda(u) \end{pmatrix} = \begin{pmatrix} F''(x_n) + \lambda_n E_{xx}(x_n,u_n) & E_x(x_n,u_n)^* \\ \\ E_x(x_n,u_n) & 0 \end{pmatrix}^{-1} RHS \tag{9.17}$$

with

$$RHS = \begin{pmatrix} F''(x_n)x_n - F'(x_n) + \lambda_n E_{xx}(x_n,u_n)x_n - E_{xu}(x_n,u_n)(u-u_n))\lambda_n \\ E_x(x_n,u_n)x_n - E_u(x_n,u_n)(u-u_n) - E(x_n,u_n) \end{pmatrix}$$

and then the second inequality of (9.16) becomes the variational inequality for $u \in \mathcal{C}$. Consequently, we obtain the following algorithm:

Sequential Programming II

1. Given $u \in \mathcal{C}$, solve (9.17) for $(x(u),\lambda(u))$.
2. Solve the variational inequality for $u \in \mathcal{C}$:

$$H(v) - H(u) + ((E_u(x(u),u)^* - E_u(x_n,u_n)^*)\lambda_n + E_u(x_n,u_n)^*\lambda(u), v-u) \geq 0, \tag{9.18}$$

   for all $v \in \mathcal{C}$.
3. Set $u_{n+1} = u$ and $x_{n+1} = x(u)$. Iterate until convergence.

Example (Optimal Control Problem) Let $(x,u) \in H^1(0,T;R^n) \times L^2(0,T;R^m)$. Consider the optimal control problem:

$$\min \quad \int_0^T (\ell(x(t)) + h(u(t)))\,dt$$

subject to the dynamical constraint

$$\frac{d}{dt}x(t) = f(x(t)) + Bu(t), \quad x(0) = x_0$$

and the control constraint

$$u \in \mathcal{C} = \{u(t) \in U, \quad a.e.\},$$

where U is a closed convex set in R^m. Then, the necessary optimality condition for (9.11) is

$$E'(x_n,u_n)(x-x_n,u-u_n) + E(x_n,u_n) = -\frac{d}{dt}x + f'(x_n)(x-x_n) + Bu = 0$$

$$F'(x) + E_x(x_n,u_n)\lambda = \frac{d}{dt}\lambda + f'(x_n)^t\lambda + \ell'(x) = 0$$

$$u(t) = \text{argmin}_{v\in U}\{h(v) + (B^t\lambda(t),v)\}.$$

If $h(u) = \frac{\alpha}{2}|u|^2$ and $U = R^m$, then $u(t) = -\frac{B^t\lambda(t)}{\alpha}$. Thus, (9.11) is equivalent to solving the two point boundary value for (x, λ):

$$\begin{cases} \dfrac{d}{dt}x = f'(x_n)(x - x_n) - \dfrac{1}{\alpha}BB^t\lambda, \quad x(0) = x_0 \\ \\ -\dfrac{d}{dt}\lambda = (f'(x) - f'(x_n))^*\lambda_n + f'(x_n)^*\lambda + \ell'(x), \quad \lambda(T) = 0. \end{cases}$$

Example (Inverse Medium problem) Let $(y, v) \in H^1(\Omega) \times H^(\Omega)$. Consider the inverse medium problem:

$$\min \quad \int_\Gamma \frac{1}{2}|y - z|^2 \, ds_x + \int_\Omega (\beta\,|v| + \frac{\alpha}{2}\,|\nabla v|^2)\, dx$$

subject to

$$-\Delta y + vy = 0 \text{ in } \Omega, \quad \frac{\partial y}{\partial n} = g \text{ at the boundary } \Gamma$$

and

$$v \in \mathcal{C} = \{0 \le v \le \gamma < \infty\},$$

where Ω is a bounded open domain in R^2, z is a boundary measurement of the voltage y and $g \in L^2(\Gamma)$ is a given current. Problem is to determine the potential function $v \in \mathcal{C}$ from $z \in L^2(\Gamma)$ in a class of sparse and clustered media. Then, the necessary optimality condition for (9.11) is given by

$$\begin{cases} -\Delta y + v_n(y - y_n) + vy_n = 0, \quad \dfrac{\partial y}{\partial n} = g \\ \\ -\Delta\lambda + v_n\lambda = 0, \quad \dfrac{\partial \lambda}{\partial n} = y - z \\ \\ \displaystyle\int_\Omega (\alpha\,(\nabla v, \nabla\phi - \nabla v) + \beta\,(|\phi| - |v|) + (\lambda y_n, v))\, dx \ge 0 \text{ for all } \phi \in \mathcal{C} \end{cases}$$

9.2 Exact penalty method

Consider the penalty method for the inequality constraint minimization

$$\min \quad F(x), \quad \text{subject to } Gx \le c \tag{9.19}$$

where $G \in \mathcal{L}(X, L^2(\Omega))$. If x^* is a minimizer of (9.19), the necessary optimality condition is given by

$$F'(x^*) + G^*\mu = 0, \quad \mu = \max(0, \mu + Gx^* - c) \text{ a.e. in } \Omega. \tag{9.20}$$

For $\beta > 0$ the penalty method is defined by

$$\min \quad F(x) + \beta\,\psi(Gx - c) \tag{9.21}$$

where

$$\psi(y) = \int_\Omega \max(0, y)\, d\omega.$$

The necessary optimality condition is given by

$$-F'(x) \in \beta G^* \partial\psi(Gx - c). \tag{9.22}$$

where

$$\partial\psi(s) = \begin{cases} 0 & s < 0 \\ [0,1] & s = 0 \\ 1 & s > 0 \end{cases}$$

Suppose the Lagrange multiplier μ satisfies

$$\sup_{\omega\in\Omega} |\mu(\omega)| \le \beta, \tag{9.23}$$

then $\mu \in \beta\, \partial\psi(Gx^* - c)$ and thus x^* satisfies (9.22). Moreover, if (9.21) has a unique minimizer in a neighborhood of x^*, then $x = x^*$ where x is a minimizer of (9.21).

Due to the singularity and the non-uniqueness of the subdifferential of $\partial\psi$, the direct treatment of the condition (9.21) many not be feasible for numerical computation. We define a regularized functional of $\max(0, s)$; for $\epsilon > 0$

$$\max{}_\epsilon(0, s) = \begin{cases} \frac{\epsilon}{2} & s \le 0 \\ \\ \frac{1}{2\epsilon}|s|^2 + \frac{\epsilon}{2} & 0 \le s \le \epsilon \\ \\ s & s \ge 0 \end{cases} .$$

and consider the regularized problem of (9.21);

$$\min \quad J_\epsilon(x) = F(x) + \beta\, \psi_\epsilon(Gx - c), \tag{9.24}$$

whee

$$\psi_\epsilon(y) = \int_\Omega \max{}_\epsilon(0, y)\, d\omega$$

Theorem (Consistency) Assume F is weakly lower semi-continuous. For an arbitrary $\beta > 0$, any weak cluster point of the solution x_ϵ, $\epsilon > 0$ of the regularized problem (9.24) converges to a solution x of the non-smooth penalized problem (9.21) as $\epsilon \to 0$. Proof: First we note that $0 \le \max_\epsilon(s) - \max(0, s) \le \dfrac{\epsilon}{2}$ for all $s \in R$. Let x be a solution to (9.21) and x_ϵ be a solution of the regularized problem (9.24). Then we have

$$F(x_\epsilon) + \beta\, \psi_\epsilon(x_\epsilon) \le F(x) + \beta\, \psi_\epsilon(x)$$

$$F(x) + \beta\, \psi(x) \le F(x_\epsilon) + \beta\, \psi(x_\epsilon).$$

Adding these inequalities,

$$\psi_\epsilon(x_\epsilon) - \psi(x_\epsilon) \le \psi_\epsilon(x) - \psi(x).$$

Thus, we have for all cluster points $\bar{x}$ of x_ϵ

$$F(\bar{x}) + \beta\, \psi(\bar{x}) \le F(x) + \beta\, \psi(x),$$

since F is weakly lower semi-continuous,

$$\psi_\epsilon(x_\epsilon) - \psi(\bar{x}) = \psi_\epsilon(x_\epsilon) - \psi(x_\epsilon) + \psi(x_\epsilon) - \psi(\bar{x})$$

$$\leq \psi_\epsilon(x) - \psi(x) + \psi(x_\epsilon) - \psi(\bar{x}).\square$$

and

$$\lim_{\epsilon\to 0^+} (\psi_\epsilon(x) - \psi(x) + \psi(x_\epsilon) - \psi(\bar{x})) \leq 0$$

As a consequence of Theorem, we have

Corollary If (9.21) has the unique minimizer in a neighborhood of x^* and (9.23) holds, then x_ϵ converges to the solution x^* of (9.19) as $\epsilon \to 0$.

The necessary optimality for (9.24) is given by

$$F'(x) + G^*\psi_\epsilon'(Gx - c) = 0. \tag{9.25}$$

Although the non-uniqueness for concerning subdifferential in the optimality condition in (9.22) is now bypassed through regularization, the optimality condition (9.25) is still nonlinear. For this objective, we first observe that $\max_\epsilon'$ has an alternative representation, i.e.,

$$\max_\epsilon'(s) = \chi_\epsilon(s)s, \quad \chi_\epsilon(s) = \begin{cases} 0 & s \leq 0 \\ \frac{1}{\max(\epsilon,s)} & s > 0 \end{cases} \tag{9.26}$$

This suggests the following semi-implicit fixed point iteration;

$$\alpha\, P(x^{k+1} - x^k) + F'(x^k) + \beta\, G^*\chi_k(Gx^k - c)(Gx^{k+1} - c)_j = 0, \quad \chi_k = \chi_\epsilon(Gx_k - c), \tag{9.27}$$

where P is positive, self-adjoint and serves a pre-conditioner for F' and $\alpha > 0$ serves a stabilizing and acceleration stepsize (see, Theorem 2).

Let $d^k = x^{k+1} - x^k$. Equation (9.29) for x^{k+1} is equivalent to the equation for d^k

$$\alpha\, Pd^k + F'(x^k) + \beta G^*\chi_k(Gx^k - c + Gd^k) = 0,$$

which gives us

$$(\alpha\, P + \beta\, G^*\chi_k G)d^k = -F_\epsilon'(x^k). \tag{9.28}$$

The direction d^k is a decent direction for $J_\epsilon(x)$ at x^k, indeed,

$$(d_k, J_\epsilon'(x^k)) = -((\alpha\, P + \beta\, G^*\chi_k G)d_k, d^k) = -\alpha(Dd^k, d^k) - \beta(\chi_k Gd^k, Gd^k) < 0.$$

Here we used the fact that P is strictly positive definite. So, the iteration (9.29) can be seen as a decent method:

Algorithm (Fixed point iteration)

Step 0. Set parameters: $\beta, \alpha, \epsilon, P$

Step 1. Compute the direction by $(\alpha\, P + \beta G^*\chi_k G)d^k = -F'(x^k)$

Step 2. Update $x^{k+1} = x^k + d^k$.

Step 3. If $|J'_\epsilon(x^k)| < TOL$, then stop. Otherwise repeat Step 1 - Step 2.

Let us make some remarks on the Algorithm. In many applications, the structure of F' and G are sparse block diagonals, and the resulting system (9.28) for the direction d^k then becomes a linear system with a sparse symmetric positive-definite matrix; and can be efficiently solved by the Cholesky decomposition method, for example. If $F(x) = \frac{1}{2}(x, Ax) - (b, x)$, then we have $F'(x) = Ax - b$. For this case we may use the alternative update

$$\alpha\, P(x^{k+1} - x^k) + Ax^{k+1} - b + \beta\, G^* \chi_k (Gx^{k+1} - c) = 0,$$

assuming that it doesn't cost much to perform this fully implicit step.

For the bilateral inequality constraint (**??**)-(**??**) the update (9.29) becomes

$$\alpha\, P(x^{k+1} - x^k) + F'(x^k) + \beta\, G(j,:)^* \chi_k (Gx^{k+1} - g)_j = 0 \tag{9.29}$$

where the active index set $\mathcal{A}_k$ and the diagonal matrix χ_k on $\mathcal{A}_k$ are defined by

$$j \in \mathcal{A}_k = \{j : (Gx^k - g)_j \geq g_j\},$$

$$(\chi_k)_{jj} = \frac{1}{\max(\epsilon, (Gx^k - g)_j)}.$$

The algorithm is globally convergent practically and the following results justify the fact.

Theorem 3 Let

$$R(x, \hat{x}) = -(F(x) - F(\hat{(x)}) - F'(\hat{x})(x - \hat{(x)})) + \alpha\, (P(x - \hat{x}), x - \hat{x}) \geq \omega\, |x - \hat{x}|^2$$

for some $\omega > 0$ and all $x,\ \hat{x}$. Then, we have

$$R(x^{k+1}, x^k) + F(x^{k+1}) - F(x^k) + \frac{\beta}{2}\,((\chi_k G d^k, G d^k)$$

$$+ \sum_{j'} (\chi_k, |Gx^{k+1} - c - g|^2 - |Gx^k - c - g|^2) + \sum_{j''} (\chi_k, |Gx^{k+1} - c + g|^2 - |Gx^k - c + g|^2),$$

where $\{j' : (Gx^k - c - g)_{j'} \geq 0\}$ and $\{j'' : (Gx^k - c + g)_{j''} \leq 0\}$.
Proof: Multiplying (9.29) by $d^k = x^{k+1} - x^k$

$$\alpha\, (Pd^k, d^k) - (F(x^{k+1}) - F(x^k) - F'(x^k) d^k)$$

$$+ F(x^{k+1}) - F(x^k) + E_k = 0,$$

where

$$E_k = \beta(\chi_k (G(j,:)x^{k+1} - c), G(j,:)d^k) = \frac{\beta}{2}\,((\chi_k G(j,:)^t d^k, G(j,:)d^k)$$

$$+ \sum_{j'} (\chi_k, |Gx^{k+1} - c - g|^2 - |Gx^k - c - g|^2) + \sum_{j''} (\chi_k, |Gx^{k+1} - c + g|^2 - |Gx^k - c + g|^2)),$$

Corollary 2 Suppose all inactive indexes remain inactive, then

$$\omega\, |x^{k+1} - x^k|^2 + J_\epsilon(x^{k+1}) - J_\epsilon(x^k) \leq 0$$

and $\{x^k\}$ is globally convergent.

Proof: Since $s \to \psi(\sqrt{|s-g|})$ on $s \geq g$ is and $s \to \psi(\sqrt{|s+g|})$ on $s \leq -g$ are concave, we have

$$(\chi_k, |(Gx^{k+1}-c-g)_{j'}|^2 - |(Gx^k-c-g)_{j'}|^2) \geq \psi_\epsilon((Gx^{k+1}-c)_{j'}) - \psi_\epsilon((Gx^k-c-g)_{j'})$$

and

$$(\chi_k, |(Gx^{k+1}-c+g)_{j''}|^2 - |(Gx^k-c-g)_{j''}|^2) \geq \psi_\epsilon((Gx^{k+1}-c)_{j''}) - \psi_\epsilon((Gx^k-c)_{j''})$$

Thus, we obtain

$$F(x^{k+1})+\beta\,\psi_\epsilon(x^{k+1})+R(x^{k+1},x^k)+\frac{\beta}{2}\,(\chi_k G(j,:)(x^{k+1}-x^k), G(j,:)(x^{k+1}-x^k)) \leq F(x^k)+\beta\,\psi_\epsilon(x^k).$$

If we assume $R(x,\hat{x}) \geq \omega\,|x-\hat{x}|^2$ for some $\omega > 0$, then $F(x^k)$ is monotonically decreasing and

$$\sum |x^{k+1}-x^k|^2 < \infty.$$

Corollary 3 Suppose $i \in \mathcal{A}_k^c = \mathcal{I}^k$ is inactive but $\psi_\epsilon(Gx_i^{k+1}-c) > 0$. Assume

$$\frac{\beta}{2}\,(\chi_k G(j,:)(x^{k+1}-x^k), G(j,:)(x^{k+1}-x^k)) - \beta \sum_{i\in\mathcal{I}^k} \psi_\epsilon(Gx_i^{k+1}-c) \geq -\omega'\,|x^{k+1}-x^k|^2$$

with $0 \leq \omega' < \omega$, then the algorithm is globally convergent.

10 Sensitivity analysis

In this section the sensitivity analysis is discussed for the parameter-dependent optimization. Consider the parametric optimization problem;

$$F(x,p) \quad x \in \mathcal{C}. \tag{10.1}$$

For given $p \in P$, a complete matrix space, let $x = x(p) \in \mathcal{C}$ is a minimizer of (10.1). For $\bar{p} \in P$ and $p = \bar{p} + t\dot{p} \in P$ with $t \in R$ and increment $\dot{p}$ we have

$$F(x(p),p) \leq F(x(\bar{p}),p), \quad F(x(\bar{p}),\bar{p}) \leq F(x(p),\bar{p}).$$

Thus,

$$F(x(p),p)-F(x(\bar{p}),\bar{p}) = F(x(p),p)-F(x(\bar{p}),p)+F(x(\bar{p}),p)-F(x(\bar{p}),\bar{p}) \leq F(x(\bar{p}),p)-F(x(\bar{p}),\bar{p})$$

$$F(x(p),p)-F(x(\bar{p}),\bar{p}) = F(x(p),\bar{p})-F(x(\bar{p}),\bar{p})+F(x(p),p)-F(x(p),\bar{p}) \geq F(x(p),p)-F(x(p),\bar{p})$$

and

$$F(x(p),p)-F(x(p),\bar{p}) \leq F(x(p),p)-F(x(\bar{p}),\bar{p}) \leq F(x(\bar{p}),p)-F(x(\bar{p}),\bar{p}).$$

Hence for $t > 0$ we have

$$\frac{F(x(\bar{p}+t\dot{p}),\bar{p}+t\dot{p})-F(x(\bar{p}+t\dot{p}),\bar{p})}{t} \leq \frac{F(x(\bar{p}+t\dot{p}),\bar{p}+t\dot{p})-F(x(\bar{p}),\bar{p})}{t} \leq \frac{F(x(\bar{p}),\bar{p}+t\dot{p})-F(x(\bar{p}),\bar{p})}{t}$$

$$\frac{F(x(\bar{p}),\bar{p})-F(x(\bar{p}),\bar{p}-t\dot{p})}{t} \leq \frac{F(x(\bar{p}),\bar{p})-F(x(\bar{p}-t\dot{p}),\bar{p}-t\dot{p})}{t} \leq \frac{F(x(\bar{p}-t\dot{p}),\bar{p})-F(x(\bar{p}-t\,\dot{p}),\bar{p}-t\dot{p})}{t}.$$

Based on this inequality we have

Theorem (Sensitivity I) Assume
(H1) there exists the continuous graph $p \to x(p) \in \mathcal{C}$ in a neighborhood of $(x(\bar{p}), \bar{p}) \in \mathcal{C} \times P$ and define the value function

$$V(p) = F(x(p), p).$$

(H2) in a neighborhood of $(x(\bar{p}), \bar{p})$

$$(x, p) \in \mathcal{C} \times Q \to F_p(x, p) \text{ is continuous.}$$

Then, the G-derivative of V at $\bar{p}$ exists and is given by

$$V'(\bar{p})\dot{p} = F_p(x(\bar{p}, \bar{p})\dot{p}.$$

Next, consider the implicit function case. Consider a functional

$$V(p) = F(x(p)), \quad E(x(p), p) = 0$$

where we assume that the constraint E is C^1 and $E(x, p) = 0$ defines a continuous solution graph $p \to x(p)$ in a neighborhood of $(x(\bar{p}), \bar{p})$.

Theorem (Sensitivity II) Assume there exists λ that satisfies the adjoint equation

$$E_x(x(\bar{p}), \bar{p})^* \lambda + F_x(x(\bar{p})) = 0$$

and assume $\epsilon_1 + \epsilon_2 = o(|p - \bar{p}|)$, where

$$\begin{aligned} \epsilon_1 &= (E(x(p), \bar{p}) - E(x(\bar{p}), \bar{p}) - E_x(x(\bar{p}), \bar{p})(x(p) - x(\bar{p})), \lambda) \\ \epsilon_2 &= F(x(p)) - F(x(\bar{p})) - F_x(x(\bar{p}))(x(p) - x(\bar{p})) \end{aligned} \tag{10.2}$$

Then, $p \to V(p)$ is differentiable at $\bar{p}$ and

$$V'(\bar{p})\dot{p} = (E_p(x(\bar{p}), \bar{p}), \lambda).$$

Proof: Note that

$$V(p) - V(\bar{p}) = F'(x(\bar{p}))(x(p) - x(\hat{p})) + \epsilon_2$$

and

$$(E_x(x(\bar{p}), \bar{p})(x(p) - x(\bar{p})), \lambda) + F_x(x(\bar{p})(x(p) - x(\bar{p})) = 0.$$

Since

$$0 = (E(x(p), p) - E(x(\bar{p}), \bar{p}), \lambda) = (E_x(x(\bar{p}), \bar{p})(x(p) - x(\bar{p})) + (E(x(p), p) - E(x(p), \bar{p}), \lambda) + \epsilon_1,$$

we have

$$V(p) - V(\bar{p}) = (E_p(x(p), \bar{p})(p - \bar{p}), \lambda) + \epsilon_1 + \epsilon_2 + o(|p - \bar{p}|)$$

and thus $V(p)$ is differentiable and the desired result. $\square$

If we assume that E and F is C^2, then it is sufficient to have the Holder continuity

$$|x(p) - x(\bar{p})|_X \sim o(|p - \bar{p}|_Q^{\frac{1}{2}})$$

for condition (10.2) holding.

Next, consider the parametric constrained optimization:

$$F(x,p) \text{ subject to } E(x,p) \in K.$$

We assume
(H3) $\lambda \in Y$ satisfies the adjoint equation

$$E_x(x(\bar{p}),\bar{p})^*\lambda + F_x(\bar{p}),\bar{p}) = 0.$$

(H4) Conditions (10.2) holds.
(H5) In a neighborhood of $(x(\bar{p}),\bar{p})$.

$$(x,p) \in \mathcal{C} \times P \to E_p(x,p) \in Y \text{ is continuous.}$$

Eigenvalue problem Let $A(p)$ be a linear operator in a Hilbert space X. For $x = (\mu, y) \in R \times X$

$$F(x) = \mu, \quad \text{subject to } A(p)y = \mu\, y.$$

That is, (μ, y) is an eigen pair of $A(p)$.

$$A(p)^*\lambda - \mu\,\lambda = 0, \quad (y,\lambda) = 1$$

Thus,

$$\mu'(\dot{p}) = (\frac{d}{dp}A(p)\dot{p}y, \lambda).$$

Next, consider the parametric constrained optimization; given $p \in Q$

$$\min_{x,u} \quad F(x,u,p) \text{ subject to } E(x,u,p) = 0. \tag{10.3}$$

Let $(\bar{x},\bar{u})$ is a minimizer given $\bar{p} \in Q$. We assume
(H3) $\lambda \in Y$ satisfies the adjoint equation

$$E_x(\bar{x},\bar{u},\bar{p})^*\lambda + F_x(\bar{x},\bar{u},\bar{p}) = 0.$$

(H4) Conditions (10.2) holds.
(H5) In a neighborhood of $(x(\bar{p}),\bar{p})$.

$$(x,p) \in \mathcal{C} \times Q \to E_p(x,p) \in Y \text{ is continuous.}$$

Theorem (Sensitivity III) Under assumptions (H1)–(H5) the necessary optimality for (12.16) is given by

$$E_x^*\lambda + F_x(\bar{x},\bar{u},\bar{p}) = 0$$

$$E_u^*\lambda + F_u(\bar{x},\bar{u},\bar{p}) = 0$$

$$E(\bar{x},\bar{u},\bar{p}) = 0$$

and

$$V'(\bar{p})\dot{p} = F_p(\bar{x},\bar{u},\bar{p})\dot{p} + (\lambda, E_p(\bar{x},\bar{u},\bar{p})\dot{p}) = 0.$$

Proof:

$$F(x(p),p)-F(x(\bar p),\bar p)=F(x(p),\bar p)-F(x(\bar p),\bar p)+F(x(p),p)-F(x(p),\bar p)$$

$$\geq F_x(x(\bar p),\bar p)(x(p)-x(\bar p)+F(x(p),p)-F(x(p),\bar p)$$

$$F(x(p),p)-F(x(\bar p),\bar p)=F(x(p),p)-F(x(\bar p),p)+F(x(\bar p),p)-F(x(\bar p),\bar p)$$

$$\leq F_x(x(p),p)(x(p)-x(\bar p))+F(x(\bar p),p)-F(x(\bar p),\bar p).$$

Note that

$$E(x(p),p)-E(x(\bar p),\bar p)=E_x(x(\bar p))(x(p)-x(\bar p)+E(x(p),p)-E(x(p),\bar p)=o(|x(p)-x(\bar p)$$

and

$$(\lambda,E_x(\bar x,\bar p)(x(p)-x(\bar p))+F_x(\bar x,\bar p),x(p)-x(\bar p)=0$$

Combining the above inequalities and letting $t\to 0$, we have

11 Augmented Lagrangian Method

In this section we discuss the augmented Lagrangian method for the constrained minimization;

$$\min \quad F(x) \text{ subject to } E(x)=0,\;\; G(x)\leq 0 \tag{11.1}$$

over $x\in\mathcal{C}$. For the equality constraint case we consider the augmented Lagrangian functional

$$L_c(x,\lambda)=F(x)+(\lambda,E(x))+\frac{c}{2}|E(x)|^2.$$

for $c>0$. The augmented Lagrangian method is an iterative method with the update

$$x_{n+1}=\text{argmin}_{x\in\mathcal{C}}\,L_c(x,\lambda_n)$$

$$\lambda_{n+1}=\lambda_n+c\,E(x_n).$$

It is a combination of the multiplier method and the penalty method. That is, if $c=0$ and $\lambda_{n+1}=\lambda_n+\alpha\ E(x_n)$ with step size $\alpha>0$ is the multiplier method and if $\lambda_n=0$ is the penalty method. The multiplier method converges (locally) if F is (locally) uniformly convex. The penalty method requires $c\to\infty$ and may suffer if $c>0$ is too large. It can be shown that the augmented Lagrangian method is locally convergent provided that $L_c''(x,\lambda)$ is uniformly positive near the solution pair $(\bar x,\bar\lambda)$. Since

$$L''(\bar x,\bar\lambda)=F''(\bar x)+(\bar\lambda,E''(\bar x))+c\,E'(\bar x)^*E'(\bar x),$$

the augmented Lagrangian method has an enlarged convergent class compared to the multiplier method. Unlike the penalty method it is not necessary to $c\to\infty$ and the Lagrange multiplier update speeds up the convergence.

For the inequality constraint case $G(x)\leq 0$ we consider the equivalent formulation:

$$\min \quad F(x)+\frac{c}{2}|G(x)-z|^2+(\mu,G(x)-z) \text{ over } (x,z)\in\mathcal{C}\times Z$$

subject to

$$G(x) = z, \quad \text{and} \quad z \le 0.$$

For minimizing this over $z \le 0$ we have that

$$z^* = \min(0, \frac{\mu + c\,G(x)}{c})$$

attains the minimum and thus we obtain

$$\min \quad F(x) + \frac{1}{2c}(|\max(0, \mu + c\,G(x))|^2 - |\mu|^2) \text{ over } x \in \mathcal{C}.$$

For (11.1), given (λ, μ) and $c > 0$, define the augmented Lagrangian functional

$$L_c(x, \lambda, \mu) = F(x) + (\lambda, E(x)) + \frac{c}{2}\,|E(x)|^2 + \frac{1}{2c}(|\max(0, \mu + c\,G(x))|^2 - |\mu|^2). \quad (11.2)$$

The first order augmented Lagrangian method is a sequential minimization of $L_c(x, \lambda, \mu)$ over $x \in \mathcal{C}$;

<u>Algorithm (First order Augmented Lagrangian method)</u>

1. Initialize (λ^0, μ^0)
2. Let x_n be a solution to
$$\min \quad L_c(x, \lambda_n, \mu_n) \text{ over } x \in \mathcal{C}.$$
3. Update the Lagrange multipliers
$$\lambda_{n+1} = \lambda_n + c\,E(x_n), \quad \mu_{n+1} = \max(0, \mu_n + c\,G(x_n)). \quad (11.3)$$

<u>Remark</u> (1) Define the value function

$$\Phi(\lambda, \mu) = \min_{x \in \mathcal{C}} \; L_c(x, \lambda, \mu).$$

Then, one can prove that

$$\Phi_\lambda = E(x), \quad \Phi_\mu = \max(0, \mu + c\,G(x))$$

(2) The Lagrange multiplier update (11.3) is a gradient method for maximizing $\Phi(\lambda, \mu)$.

(3) For the equality constraint case

$$L_c(x, \lambda) = F(x) + (\lambda, E(x)) + \frac{c}{2}|E(x)|^2$$

and

$$L_c'(x, \lambda) = L_0'(x, \lambda + c\,E(x)).$$

The necessary optimality for

$$\max_\lambda \min_x \quad L_c(x, \lambda)$$

is given by

$$L_c'(x, \lambda) = L_0'(x, \lambda + c\,E(x)) = 0, \quad E(x) = 0 \quad (11.4)$$

If we apply the Newton method to system (11.4), we obtain

$$\begin{pmatrix} L_0''(x, \lambda + c\,E(x)) + c\,E'(x)^*E'(x) & E'(x)^* \\ \\ E'(x) & 0 \end{pmatrix} \begin{pmatrix} x^+ - x \\ \\ \lambda^+ - \lambda \end{pmatrix} = - \begin{pmatrix} L_0'(x, \lambda + c\,E(x)) \\ \\ E(x) \end{pmatrix}.$$

The $c > 0$ adds the coercivity on the Hessian L_0'' term.

11.1 Primal-Dual Active set method

For the inequality constraint $G(x) \leq 0$ ($e.g., Gx - \tilde{c} \leq 0$);

$$L_c(x,\mu) = F(x) + \frac{1}{2c}(|\max(0, \mu + c\,G(x)|^2 - |\mu|^2)$$

and the necessary optimality condition is

$$F'(x) + G'(x)^*\mu, \quad \mu = \max(0, \mu + c\,G(x))$$

where $c > 0$ is arbitrary. Based on the complementarity condition we have primal dual active set method.

<u>Primal Dual Active Set method</u>

For the affine case $G(x) = Gx - \tilde{c}$ we have the Primal Dual Active Set method as:

1. Define the active index and inactive index by

$$\mathcal{A} = \{k \in (\mu + c\,(Gx - \tilde{c}))_k > 0\}, \quad \mathcal{I} = \{j \in (\mu + c\,(Gx - \tilde{c}))_j \leq 0\}$$

2. Let $\mu^+ = 0$ on $\mathcal{I}$ and $Gx^+ = \tilde{c}$ on $\mathcal{A}$.
3. Solve for $(x^+, \mu^+_{\mathcal{A}})$

$$F''(x)(x^+ - x) + F'(x) + G^*_{\mathcal{A}}\mu^+ = 0, \quad G_{\mathcal{A}}x^+ - c = 0$$

4. Stop or set $n = n + 1$ and return to step 1,

where $G_{\mathcal{A}} = \{G_k\},\ k \in \mathcal{A}$. For the nonlinear G

$$G'(x)(x^+ - x) + G(x) = 0 \text{ on } \mathcal{A} = \{k \in (\mu + c\,G(x))_k > 0\}.$$

The primal dual active set method is a semismooth Newton method. That is, if we define a generalized derivative of $s \to \max(0, s)$ by 0 on $(-\infty, 0]$ and 1 on $(0, \infty)$. Note that $s \to \max(0, s)$ is not differentiable at $s = 0$ and we define the derivative at $s = 0$ as the limit of derivative for $s < 0$. So, we select the generalized derivative of $\max(0, s)$ as $0,\ s \leq 0$ and $1,\ s > 0$ and thus the generalized Newton update

$$\begin{array}{ll} \mu^+ - \mu + 0 = \mu & \text{if } \mu + c\,(Gx - c) \leq 0 \\ \\ c\,G(x^+ - x) + c\,(Gx - \tilde{c}) = Gx^+ - \tilde{c} = 0 & \text{if } \mu + c\,(Gx - \tilde{c}) > 0, \end{array}$$

which results in the active set strategy

$$\mu^+_j = 0,\ \ j \in \mathcal{I} \quad \text{and} \quad (Gx^+ - \tilde{c})_k = 0,\ \ k \in \mathcal{A}.$$

We will introduce the class of semismooth functions and the semismooth Newton method in Chapter 7 in function spaces. Specifically, the pointwise (coordinate) operation $s \to \max(0, s)$ defines a semismooth function from $L^p(\Omega) \to L^q(\Omega),\ p > q$.

12 Lagrange multiplier Theory for Nonsmooth convex Optimization

In this section we present the Lagrange multiplier theory for the nonsmmoth optimization of the form

$$\min \quad f(x)+\varphi(\Lambda x) \text{ over } x\in\mathcal{C} \tag{12.1}$$

where X is a Banach space, H is a Hilbert space lattice, $f : X \to R$ is C^1, $\Lambda \in \mathcal{L}(X, H)$ and $\mathcal{C}$ is a closed convex set in X. The nonsmoothness is represented by φ and φ is a proper, lower semi continuous convex functional in H.

This problem class encompasses a wide variety of optimization problems including variational inequalities of the first and second kind [**?**, **?**]. We will specify those to our specific examples. Moreover, the exact penalty formulation of the constraint problem is

$$\min \quad f(x)+c\,|E(x)|_{L^1}+c\,|(G(x))^+|_{L^1} \tag{12.2}$$

where $(G(x))^+ = \max(0, G(x))$.

We describe the Lagrange multiplier theory to deal with the non-smoothness of φ. Note that (12.1) is equivalent to

$$\begin{aligned} &\min \quad f(x)+\varphi(\Lambda x-u) \\ &\text{subject to } x\in\mathcal{C} \quad \text{and} \quad u=0 \text{ in } H. \end{aligned} \tag{12.3}$$

Treating the equality constraint $u = 0$ in 12.3) by the augmented Lagrangian method, results in the minimization problem

$$\min_{x\in C, u\in H} \quad f(x)+\varphi(\Lambda x-u)+(\lambda,\,u)_H+\frac{c}{2}\,|u|_H^2, \tag{12.4}$$

where $\lambda \in H$ is a multiplier and c is a positive scalar penalty parameter. By (12.6) we have

$$\min_{y\in C} \quad L_c(y,\lambda)=f(x)+\varphi_c(\Lambda x,\lambda). \tag{12.5}$$

where

$$\begin{aligned} \varphi_c(u,\lambda) &= \inf\nolimits_{v\in H}\,\{\varphi(u-v)+(\lambda,\,v)_H+\tfrac{c}{2}\,|v|_H^2\} \\ &= \inf_{z\in H}\,\{\varphi(z)+(\lambda,u-z)_H+\frac{c}{2}\,|u-z|_H^2\} \;\; (u-v=z). \end{aligned} \tag{12.6}$$

For $u,\ \lambda \in H$ and $c > 0$, $\varphi_c(u, \lambda)$ is called the generalized Yoshida-Moreau approximations φ. It can be shown that $u \to \varphi_c(u, \lambda)$ is continuously Fréchet differentiable with Lipschitz continuous derivative. Then, the augmented Lagrangian functional [**?**, **?**, **?**] of (12.1) is given by

$$L_c(x,\lambda)=f(x)+\varphi_c(\Lambda x,\lambda). \tag{12.7}$$

Let φ^* is the convex conjugate of φ

$$\begin{aligned} \varphi_c(u,\lambda) &= \inf\nolimits_{z\in H}\{\sup\nolimits_{y\in H}\{(y,z)-\varphi^*(y)\}+(\lambda,u-z)_H+\tfrac{c}{2}\,|u-z|_H^2\} \\ &= \sup_{y\in H}\{-\frac{1}{2c}|y-\lambda|^2+(y,u)-\varphi^*(y)\}. \end{aligned} \tag{12.8}$$

where we used

$$\inf_{z\in H}\{(y,z)_H+(\lambda,u-z)_H+\frac{c}{2}\,|u-z|_H^2\}=(y,u)-\frac{1}{2c}|y-\lambda|_H^2$$

Thus,

$$\varphi_c'(u,\lambda)=y_c(u,\lambda)=\mathrm{argmax}_{y\in H}\{-\frac{1}{2c}|y-\lambda|^2+(y,u)-\varphi^*(y)\}=\lambda+c\,v_c(u,\lambda)\quad(12.9)$$

where

$$v_c(u,\lambda)=\mathrm{argmin}_{v\in H}\{\varphi^*(u-v)+(\lambda,v)+\frac{c}{2}|v|^2\}.$$

Moreover, if

$$p_c(\lambda)=\mathrm{argmin}\{\frac{1}{2c}|p-\lambda|^2+h^*(p)\},$$

then

$$\varphi_c'(u,\lambda)=p_c(\lambda+c\,x).$$

Theorem (Lipschitz complementarity)
(1) If $\lambda\in\partial\varphi(x)$ for $x,\ \lambda\in H$, then $\lambda=\varphi_c'(x,\lambda)$ for all $c>0$.
(2) Conversely, if $\lambda=\varphi_c'(x,\lambda)$ for some $c>0$, then $\lambda\in\partial\varphi(x)$.

Proof: If $\lambda\in\partial\varphi(x)$, then from (12.6) and (12.8)

$$\varphi(x)\geq\varphi_c(x,\lambda)\geq\langle\lambda,x\rangle-\varphi^*(\lambda)=\varphi(x).$$

Thus, $\lambda\in H$ attains the supremum of (12.17) and we have $\lambda=\varphi_c'(x,\lambda)$. Conversely, if $\lambda\in H$ satisfies $\lambda=\varphi_c'(x,\lambda)$ for some $c>0$, then $v_c(x,\lambda)=0$ by (12.17). Hence it follows from that

$$\varphi(x)=\varphi_c(x,\lambda)=(\lambda,x)-\varphi^*(\lambda).$$

which implies $\lambda\in\partial\varphi(x)$. $\square$

If $x_c\in\mathcal{C}$ denotes the solution to (12.5), then it satisfies

$$\begin{aligned}&\langle f'(x_c)+\Lambda^*\lambda_c,x-x_c\rangle_{X^*,X}\geq 0,\quad\text{for all }x\in\mathcal{C}\\&\lambda_c=\varphi_c'(\Lambda x_c,\lambda).\end{aligned}\qquad(12.10)$$

Under appropriate conditions [**?**, **?**, **?**] the pair $(x_c,\lambda_c)\in\mathcal{C}\times H$ has a (strong-weak) cluster point $(\bar{x},\bar{\lambda})$ as $c\to\infty$ such that $\bar{x}\in\mathcal{C}$ is the minimizer of (12.1) and that $\bar{\lambda}\in H$ is a Lagrange multiplier in the sense that

$$\langle f'(\bar{x})+\Lambda^*\bar{\lambda},x-\bar{x}\rangle_{X^*,X}\geq 0,\quad\text{for all }x\in\mathcal{C},\qquad(12.11)$$

with the complementarity condition

$$\bar{\lambda}=\varphi_c'(\Lambda\bar{x},\bar{\lambda}),\quad\text{for each }c>0.\qquad(12.12)$$

System (12.11)–(12.12) is for the primal-dual variable $(\bar{x},\bar{\lambda})$. The advantage here is that the frequently employed differential inclusion $\bar{\lambda}\in\partial\varphi(\Lambda\bar{x})$ is replaced by the equivalent nonlinear equation (12.12).

The first order augmented Lagrangian method for (12.1) is given by

- Select λ^0 and set $n = 0$
- Let $x_n = \text{argmin}_{x\in\mathcal{C}} L_c(x, \lambda_n)$
- Update the Lagrange multiplier by $\lambda_{n+1} = \varphi_c'(\Lambda x_n, \lambda_n)$.
- Stop or set $n = n + 1$ and return to step 1.

In many applications, the convex conjugate functional φ^* of φ is given by

$$\varphi^*(v) = I_{K^*}(v),$$

where K^* is a closed convex set in H and I_S is the indicator function of a set S. From (**??**)

$$\varphi_c(u, \lambda) = \sup_{y\in K^*} \{-\frac{1}{2c}|y - \lambda|^2 + (y, u)\}$$

and (12.12) is equivalent to

$$\bar{\lambda} = Proj_{K^*}(\bar{\lambda} + c\,\Lambda\bar{x}). \tag{12.13}$$

which is the basis of our approach.

12.1 Examples

In this section we discuss the applications of the Lagrange multiplier theorem and the complementarity.
Example (Inequality) Let $H = L^2(\Omega)$ and $\varphi = I_{\mathcal{C}}$ with $\mathcal{C} = \{s \in L^2(\Omega) : z \le 0,\ a.e.$, i.e.,

$$\min \quad f(x) \text{ over } x \in \Lambda x - c \le 0. \tag{12.14}$$

In this case

$$\varphi^*(v) = I_{K^*}(v) \text{ with } K^* = -\mathcal{C}$$

and the complementarity (15.25) implies that

$$f'(\bar{x}) + \Lambda^*\bar{\lambda} = 0$$

$$\bar{\lambda} = \max(0, \bar{\lambda} + c\,(\Lambda\bar{x} - c)).$$

Example ($L^1(\Omega)$ optimization) Let $H = L^2(\Omega)$ and $\varphi(v) = \displaystyle\int_\Omega |v|\,dx$, i.e.

$$\min \quad f(y) + \int_\Omega |\Lambda y|_2\,dx. \tag{12.15}$$

In this case

$$\varphi^*(v) = I_{K^*}(v) \text{ with } K^* = \{z \in L^2(\Omega) : |z|_2 \le 1,\ a.e.\}$$

and the complementarity (15.25) implies that

$$f'(\bar{x}) + \Lambda^*\bar{\lambda} = 0$$

$$\bar{\lambda} = \frac{\bar{\lambda} + c\,\Lambda\bar{x}}{\max(1, |\bar{\lambda} + c\,\Lambda\bar{x}|_2)} \quad \text{a.e..}$$

Exaples (Constrained optimization) Let X be a Hilbert space and consider

$$\min \quad f(x) \text{ subject to } \Lambda x \in \mathcal{C} \tag{12.16}$$

where $f : X \to R$ is C^1 and $\mathcal{C}$ is a closed convex set in X. In this case we let $f = J$ and Λ is the natural injection and $\varphi = I_{\mathcal{C}}$.

$$\varphi_c(u,\lambda) = \inf_{z \in \mathcal{C}} \{(\lambda, u - z)_H + \frac{c}{2}\,|u - z|_H^2\}. \tag{12.17}$$

and

$$z^* = Proj_{\mathcal{C}}(u + \frac{\lambda}{c}) \tag{12.18}$$

attains the minimum in (12.17) and

$$\varphi_c'(u,\lambda) = \lambda + c\,(u - z^*) \tag{12.19}$$

If $H = L^2(\Omega)$ and $\mathcal{C}$ is the bilateral constraint $\{y \in X : \phi \le \Lambda y \le \psi\}$, the complementarity (12.18)–(12.19) implies that for $c > 0$

$$\begin{aligned} &f'(x^*) + \Lambda^* \mu = 0 \\ &\mu = \max(0, \mu + c\,(\Lambda x^* - \psi) + \min(0, \mu + c\,(\Lambda x^* - \phi)) \qquad \text{a.e..} \end{aligned} \tag{12.20}$$

If either $\phi = -\infty$ or $\psi = \infty$, it is unilateral constrained and defines the generalized obstacle problem. The cost functional $J(y)$ can represent the performance index for the optimal control and design problem, the fidelity of data-to-fit for the inverse problem and deformation and restoring energy for the variational problem.

13 Semismooth Newton method

In this section we present the semismooth Newton method for the nonsmooth necessary optimality condition, e.g., the complementarity condition (12.10):

$$\lambda = \varphi_c'(\Lambda x, \lambda).$$

Consider the nonlinear equation $F(y) = 0$ in a Banach space X. The generalized Newton update is given by

$$y^{k+1} = y^k - V_k^{-1} F(y^k), \tag{13.1}$$

where V_k is a generalized derivative of F at y^k. In the finite dimensional space or for a locally Lipschitz continuous function F let D_F denote the set of points at which F is differentiable. For $x \in X = R^n$ we define the Bouligand derivative $\partial_B F(x)$ as

$$\partial_B F(x) = \left\{ J : J = \lim_{x_i \to x,\ x_i \in D_F} \nabla F(x_i) \right\}, \tag{13.2}$$

where D_F is dense by Rademacher's theorem which states that every locally Lipschitz continuous function in the finite dimensional space is differentiable almost everywhere. Thus, we take $V_k \in \partial_B F(y^k)$.

In infinite dimensional spaces notions of generalized derivatives for functions which are not C^1 cannot rely on Rademacher's theorem. Here, instead, we shall mainly utilize a concept of generalized derivative that is sufficient to guarantee superlinear convergence of Newton's method [?]. This notion of differentiability is called Newton derivative and is defined below. We refer to [?, ?, ?, ?] for further discussion of the notions and topics. Let X, Z be real Banach spaces and let $D \subset X$ be an open set.

Definition (Newton differentiable) (1) $F\colon D \subset X \to Z$ is called Newton differentiable at x, if there exists an open neighborhood $N(x) \subset D$ and mappings $G\colon N(x) \to \mathcal{L}(X,Z)$ such that

$$\lim_{|h|\to 0} \frac{|F(x+h) - F(x) - G(x+h)h|_Z}{|h|_X} = 0.$$

The family $\{G(y) : y \in N(x)\}$ is called a N-derivative of F at x.

(2) F is called semismooth at y, if it is Newton differentiable at y and

$$\lim_{t\to 0^+} G(y + t\,h)h \quad \text{exists uniformly in } |h| = 1.$$

Semi-smoothness was originally introduced in [?] for scalar-valued functions. Convex functions and real-valued C^1 functions are examples for such semismooth functions [?, ?] in the finite dimensional space.

For example, if $F(y)(s) = \psi(y(s))$, point-wise, then $G(y)(s) \in \psi_B(y(s))$ is an N-derivative in $L^p(\Omega) \to L^q(\Omega)$ under appropriate conditions [?]. We often use $\psi(s) = |s|$ and $\max(0,s)$ for the necessary optimality.

Suppose $F(y^*) = 0$, Then, $y^* = y^k + (y^* - y^k)$ and

$$|y^{k+1} - y^*| = |V_k^{-1}(F(y^*) - F(y^k) - V_k(y^* - y^k)| \le |V_k^{-1}|o(|y^k - y^*|).$$

Thus, the semismooth Newton method is q-superlinear convergent provided that the Jacobian sequence V_k is uniformly invertible as $y^k \to y^*$. That is, if one can select a sequence of quasi-Jacobian V_k that is consistent, i.e., $|V_k - G(y^*)|$ as $y^k \to y^*$ and invertible, (13.1) is still q-superlinear convergent.

13.1 Bilateral constraint

Consider the bilateral constraint problem

$$\min \quad f(x) \text{ subject to } \phi \le \Lambda x \le \psi. \tag{13.3}$$

We have the optimality condition (12.20):

$$f'(x) + \Lambda^*\mu = 0$$

$$\mu = \max(0, \mu + c\,(\Lambda x - \psi) + \min(0, \mu + c\,(\Lambda x - \phi)) \quad \text{a.e..}$$

Since

$$\partial_B \max(0,s) = \{0,\ 1\}, \quad \partial_B \min(0,s) = \{-1,\ 0\}, \text{ at } s = 0,$$

the semismooth Newton method for the bilateral constraint (13.3) is of the form of the primal-dual active set method:

Primal dual active set method

1. Initialize $y^0 \in X$ and $\lambda^0 \in H$. Set $k = 0$.
2. Set the active set $\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^-$ and the inactive set $\mathcal{I}$ by

$$\mathcal{A}^+ = \{x \in \Omega : \mu^k + c\,(\Lambda y^k - \psi) > 0\}, \quad \mathcal{A}^- = \{x \in \Omega : \mu^k + c\,(\Lambda y^k - \phi) < 0\}, \quad \mathcal{I} = \Omega \backslash \mathcal{A}$$

3. Solve for (y^{k+1}, λ^{k+1})

$$J''(y^k)(y^{k+1} - y^k) + J'(y^k) + \Lambda^* \mu^{k+1} = 0$$

$$\Lambda y^{k+1}(x) = \psi(x), \;\; x \in \mathcal{A}^+, \;\; \Lambda y^{k-1}(x) = \phi(x), \;\; x \in \mathcal{A}^-, \;\; \mu^{k+1}(x) = 0, \;\; x \in \mathcal{I}$$

4. Stop, or set $k = k + 1$ and return to the second seep.

Thus the algorithm involves solving the linear system of the form

$$\begin{pmatrix} A & \Lambda^*_{\mathcal{A}^+} & \Lambda^*_{\mathcal{A}^-} \\ \Lambda_{\mathcal{A}^+} & 0 & 0 \\ \Lambda_{\mathcal{A}^-} & 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ \mu_{\mathcal{A}^+} \\ \mu_{\mathcal{A}^-} \end{pmatrix} = \begin{pmatrix} f \\ \psi \\ \phi \end{pmatrix}.$$

In order to gain the stability the following one-parameter family of the regularization can be used [**?**, **?**]

$$\mu = \alpha\,(\max(0, \mu + c\,(\Lambda y - \psi)) + \min(0, \mu + c\,(\Lambda y - \phi))), \quad \alpha \to 1^+.$$

13.2 L^1 optimization

As an important example, we discuss the case of $\varphi(v) = \int_\Omega |v|_2\,dx$ (L^1-type optimization). The complementarity is reduced to the form

$$\lambda \max(1 + \epsilon, |\lambda + c\,v|) = \lambda + c\,v \tag{13.4}$$

for $\epsilon \geq 0$. It is often convenient (but not essential) to use very small $\epsilon > 0$ to avoid the singularity for the implementation of algorithm. Let

$$\varphi_\epsilon(s) = \begin{cases} \frac{s^2}{2\epsilon} + \frac{\epsilon}{2}, & \text{if } |s| \leq \epsilon, \\ \\ |s|, & \text{if } |s| \geq \epsilon. \end{cases} \tag{13.5}$$

Then, (13.4) corresponds to the regularized one of (4.24):

$$\min \quad J_\epsilon(y) = F(y) + \varphi_\epsilon(\Lambda y).$$

The semismooth Newton method is given by

$$\begin{aligned} &F'(y^+) + \Lambda^* \lambda^+ = 0 \\ &\begin{cases} \lambda^+ = \dfrac{v^+}{\epsilon} & \text{if } |\lambda + v| \leq 1 + \epsilon, \\ |\lambda + c\,v|\,\lambda^+ + \left(\lambda \left(\dfrac{\lambda + c\,v}{|\lambda + c\,v|}\right)^t\right)(\lambda^+ + c\,v^+) & \\ \qquad\qquad = \lambda^+ + c\,v^+ + |\lambda + c\,v|\,\lambda & \text{if } |\lambda + c\,v| > 1 + \epsilon. \end{cases} \end{aligned} \tag{13.6}$$

There is no guarantee that (13.6) is solvable for λ^+ and is stable. In order to obtain the compact and unconditionally stable formula we use the damped and regularized algorithm with $\beta \le 1$;

$$\begin{cases} \lambda^+ = \dfrac{v}{\epsilon} & \text{if } |\lambda + c\,v| \le 1+\epsilon, \\[2ex] |\lambda + c\,v|\,\lambda^+ - \beta \left(\dfrac{\lambda}{\max(1,|\lambda|)} \left(\dfrac{\lambda + c\,v}{|\lambda + c\,v|} \right)^t \right) (\lambda^+ + c\,v^+) & \\ \qquad\qquad = \lambda^+ + c\,v^+ + \beta\,|\lambda + c\,v|\,\dfrac{\lambda}{\max(1,|\lambda|)} & \text{if } |\lambda + c\,v| > 1+\epsilon. \end{cases} \tag{13.7}$$

Here, the purpose of the regularization $\frac{\lambda}{|\lambda| \wedge 1}$ is to automatically constrain the dual variable λ into the unit ball. The damping factor β is automatically selected to achieve the stability. Let

$$d = |\lambda + c\,v|, \quad \eta = d-1, \quad a = \frac{\lambda}{|\lambda| \wedge 1}, \quad b = \frac{\lambda + c\,v}{|\lambda + c\,v|}, \quad F = ab^t.$$

Then, (13.7) is equivalent to

$$\lambda^+ = (\eta\,I + \beta\,F)^{-1}((I - \beta\,F)(c\,v^+) + \beta d\,a),$$

where by Sherman–Morrison formula

$$(\eta\,I + \beta\,F)^{-1} = \frac{1}{\eta}\,(I - \frac{\beta}{\eta + \beta\,a \cdot b} F).$$

Then,

$$(\eta\,I + \beta\,F)^{-1} \beta d\,a = \frac{\beta\,d}{\eta + \beta\,a \cdot b}\,a.$$

Since $F^2 = (a \cdot b)\,F$,

$$(\eta\,I + \beta\,F)^{-1}(I - \beta\,F) = \frac{1}{\eta}(I - \frac{\beta d}{\eta + \beta\,a \cdot b}\,F).$$

In order to achieve the stability, we let

$$\frac{\beta\,d}{\eta + \beta\,a \cdot b} = 1, \quad i.e., \quad \beta = \frac{d-1}{d - a \cdot b} \le 1.$$

Consequently, we obtain a compact Newton step

$$\lambda^+ = \frac{1}{d-1}(I - F)(c\,v^+) + \frac{\lambda}{|\lambda| \wedge 1}, \tag{13.8}$$

which results in;

<u>Primal-dual active set method (L^1-optimization)</u>

1. Initialize: $\lambda^0 = 0$ and solve $F'(y^0) = 0$ for y^0. Set $k = 0$.
2. Set inactive set $\mathcal{I}_k$ and active set $\mathcal{A}_k$ by

$$\mathcal{I}_k = \{|\lambda^k + c\,\Lambda y^k| > 1+\epsilon\}, \quad \text{and} \quad \mathcal{A}_k = \{|\lambda^k + c\,\Lambda y^k| \le 1+\epsilon\}.$$

3. Solve for $(y^{k+1},\lambda^{k+1})\in X\times H$:

$$\begin{cases} F'(y^{k+1})+\Lambda^*\lambda^{k+1}=0\\ \\ \lambda^{k+1}=\dfrac{1}{d^k-1}(I-F^k)(c\,\Lambda y^{k+1})+\dfrac{\lambda^k}{|\lambda^k|\wedge 1}\text{ in }\mathcal{A}_k\quad\text{and}\\ \\ \Lambda y^{k+1}=\epsilon\,\lambda^{k+1}\text{ in }\mathcal{I}_k.\end{cases}$$

4. Convergent or set $k=k+1$ and Return to Step 2.

This algorithm is unconditionally stable and is rapidly convergent for our test examples.
Remark (1) Note that

$$(\lambda^+,v^+)=\frac{|v^+|^2-(a\cdot v^+)(b\cdot v^+)}{d-1}+a\cdot v^+,\tag{13.9}$$

which implies stability of the algorithm. In fact, since

$$\begin{cases} F'(y^{k+1})-F'(y^k)+\Lambda^*(\lambda^{k+1}-\lambda^k)\\ \\ \lambda^{k+1}-\lambda^k=(I-F^k)(c\,\Lambda y^{k+1})+(\dfrac{1}{|\lambda^k|}-1)\,\lambda^k,\end{cases}$$

we have

$$(F'(y^{k+1})-F'(y^k),y^{k+1}-y^k)+c\,((I-F^k)\Lambda y^{k+1},\Lambda y^{k+1})+(\frac{1}{|\lambda^k|}-1)\,(\lambda^k,\Lambda y^{k+1}).$$

Suppose

$$F(y)-F(x)-(F'(y),y-x)\ge\omega\,|y-x|_X^2,$$

we have

$$F(y^k)+\sum_{j=1}^k\omega\,|y^j-y^{j-1}|^2+(\frac{1}{|\lambda^{j-1}|\wedge 1}-1)\,(\lambda^{j-1},v^j)\le F(y^0).$$

where $v^j=\Lambda y^j$. Note that $(\lambda^k,v^{k+1})>0$ implies that $(\lambda^{k+1},v^{k+1})>0$ and thus inactive at $k{+}1$-step, pointwise. This fact is a key step for proving a global convergence of the algorithm.
(2) If $a\cdot b\to 1^+$, then $\beta\to 1$. Suppose (λ,v) is a fixed point of (13.8), then

$$\left(1-\frac{1}{|\lambda|\wedge 1}\left(1-\frac{c}{d-1}\frac{\lambda+c\,v}{|\lambda+c\,v|}\cdot v\right)\right)\lambda=\frac{1}{d-1}v.$$

Thus, the angle between λ and $\lambda+c\,v$ is zero and

$$1-\frac{1}{|\lambda|\wedge 1}=\frac{c}{|\lambda|+c\,|v|-1}\left(\frac{|v|}{|\lambda|}-\frac{|v|}{|\lambda|\wedge 1}\right),$$

which implies $|\lambda|=1$. It follows that

$$\lambda+c\,v=\frac{\lambda+c\,v}{\max(|\lambda+c\,v|,1+\epsilon)}.$$

That is, if the algorithm converges, $a \cdot b \to 1$ and $|\lambda| \to 1$ and it is consistent.
(3) Consider the substitution iterate:

$$\begin{cases} F'(y^+) + \Lambda^* \lambda^+ = 0, \\ \\ \lambda^+ = \dfrac{1}{\max(\epsilon, |v|)} v^+, \quad v = \Lambda y. \end{cases} \tag{13.10}$$

Note that

$$\left(\frac{v}{|v|}, v^+ - v\right) = \left(\frac{|v^+|^2 - |v|^2 + |v^+ - v|^2}{2|v|}\right) dx$$

$$= \left(|v^+| - |v| - \frac{(|v^+| - |v|)^2 + |v^+ - v|^2}{2|v|}\right).$$

Thus,

$$J_\epsilon(v^+) \leq J_\epsilon(v),$$

where the equality holds only if v^+. This fact can be used to prove the iterative method (13.10) is globally convergent [**?**, **?**]. It also suggests to use the hybrid method; for $0 < \mu < 1$

$$\lambda^+ = \frac{\mu}{\max(\epsilon, |v|)} v^+ + (1 - \mu) \left(\frac{1}{d-1}(I - F)v^+ + \frac{\lambda}{\max(|\lambda|, 1)}\right),$$

in order to gain the global convergence property without loosing the fast convergence of the Newton method.

14 Exercise

Problem 1 If A is a symmetric matrix on R^n and $b \in R^n$ is a vector. Let $F(x) = \frac{1}{2} x^t A x - b^t x$. Show that $F'(x) = Ax - b$.

Problem 2 Sequences $\{f_n\}$, $\{g_n\}$, $\{h_n\}$ are Cauchy in $L^2(0,1)$ but not on $C[0,1]$.
Problem 3 Show that $||x| - |y|| \leq |x - y|$ and thus the norm is continuous.
Problem 4 Show that all norms are equivalent on a finite dimensional vector space.

Problem 5 Show that for a closed linear operator T, the domain $\mathcal{D}(T)$ is a Banach space if it is equipped by the graph norm

$$|x|_{\mathcal{D}(T)} = (|x|_X^2 + |Tx|_Y^2)^{1/2}.$$

Problem 6 Complete the DC motor problem.
Problem 7 Consider the spline interpolation problem (3.8). Find the Riesz representations F_i and $\tilde{F}_j$ in $H_0^2(0,1)$. Complete the spline interpolation problem (3.8).

Problem 8 Derive the necessary optimality for L^1 optimization (3.32) with $U = [-1, 1]$.

Problem 9 Solve the constrained minimization in a Hilbert space X;

$$\min \quad \frac{1}{2}|x|^2 - (a,x)$$

subject to

$$(y_i, x) = b_i,\ 1 \le i \le m_1, \quad (\tilde{y}_j, x) \le c_j,\ 1 \le j \le m_2.$$

over $x \in X$. We assume $\{y_i\}_{i=1}^{m_1},\ \{\tilde{y}_j\}_{j=1}^{m_2}$ are linearly independent.

Problem 10 Solve the pointwise obstacle problem:

$$\min \quad \int_0^1 (\frac{1}{2}|\frac{du}{dx}|^2 - f(x)u(x))\,dx$$

subject to

$$u(\frac{1}{3}) \le c_1, \quad u(\frac{2}{3}) \le c_2,$$

over a Hilbert space $X = H_0^1(0,1)$ for $f = 1$.

Problem 11 Find the differential form of the necessary optimality for

$$\min \quad \int_0^1 (\frac{1}{2}(a(x)|\frac{du}{dx}|^2 + c(x)|u|^2) - f(x)u(x))\,dx + \frac{1}{2}(\alpha_1|u(0)|^2 + \alpha_2|u(1)|^2) - (g_1, u(0)) - (g_2, u(1))$$

over a Hilbert space $X = H^1(0,1)$.

Problem 12 Let $B \in R^{m,n}$. Show that $R(B) = R^m$ if and only if BB^t is positive definite on R^m. Show that $x^* = B^t(BB^t)^{-1}c$ defines a minimum norm solution to $Bx = c$.

Problem 13 Consider the optimal control problem (3.44) (Example Control problem) with $\hat{U} = \{u \in R^m : |u|_\infty \le 1\}$. Derive the optimality condition (3.45).

Problem 14 Consider the optimal control problem (3.44) (Example Control problem) with $\hat{U} = \{u \in R^m : |u|_2 \le 1\}$. Find the orthogonal projection $P_{\mathcal{C}}$ and derive the necessary optimality and develop the gradient method.

Problem 15 Let $J(x) = \int_0^1 |x(t)|\,dt$ is a functional on $C(0,1)$. Show that $J'(d) = \int_0^1 \text{sign}(x(t))d(t)\,dt$.

Problem 16 Develop the Newton methods for the optimal control problem (3.44) without the constraint on u (i.e., $\hat{U} = R^m$).

Problem 17 (1) Show that the conjugate functional

$$h^*(p) = \sup_u\{(p,u) - h(u)\}$$

is convex.
(2) The set-valued function Φ

$$\Phi(p) = \text{argmax}_u\{(p,u) - h(u)\}$$

is monotone.

Problem 18 Find the graph $\Phi(p)$ for

$$\max_{|u|\leq\gamma}\{pu-|u|_0\}$$

and

$$\max_{|u|\leq\gamma}\{pu-|u|_1\}.$$

15 Non-convex nonsmooth optimization

In this section we consider a general class of nonsmooth optimization problems. Let X be a Banach space and continuously embed to $H=L^2(\Omega)^m$. Let $J:X\to R^+$ is C^1 and N is nonsmooth functional on H is given by

$$N(y)=\int_\Omega h(y(\omega))\,d\omega.$$

where $h:R^m\to R$ is lower semi-continuous. Consider the minimization on X;

$$\min \quad J(x)+N(x) \tag{15.1}$$

subject to

$$x\in\mathcal{C}=\{x(\omega)\in U \text{ a.e. }\}$$

where U is a closed convex set in R^m.

For example, we consider $h(u)=\frac{\alpha}{2}|u|^2+|u|_0$ where $|s|_0=1,\ s\neq 0$ and $|0|_0=0$. For an integrable function u

$$\int_\Omega |u(\omega)|_0\,d\omega=\text{meas}(\{u\neq 0\}).$$

Thus, one can formulate the volume control and scheduling problem as (15.1).

15.1 Existence

In this section we develop a existence method for a general class of minimizations

$$\min G(x) \tag{15.2}$$

on a Banach space X, including (15.1). Here, the constrained problem $x\in\mathcal{C}$ is formulated as $G(x)=G(x)+I_{\mathcal{C}}(x)$ using the indicator function of the constrained set $\mathcal{C}$. Let X be a reflexible space or $X=\tilde{X}^*$. Thus $x_n\rightharpoonup x$ means the weak or weakly star convergence in X. To prove the existence of solutions to (15.2) we modify a general existence result given in [**?**].

Theorem (Existence) Let G be an extended real-valued functional on a Banach space X satisfying the following properties:

(i) G is proper, weakly lower semi-continuous and $G(x)+\epsilon\,|x|^p$ for each $\epsilon>0$ is coercive for some p.

(ii) for any sequence $\{x_n\}$ in X with $|x_n| \to \infty$, $\frac{x_n}{|x_n|} \rightharpoonup \bar{x}$ weakly in X, and $\{G(x_n)\}$ bounded from above, we have $\frac{x_n}{|x_n|} \to \bar{x}$ strongly in X and there exists $\rho_n \in (0, |x_n|]$ and $n_0 = n_0(\{x_n\})$ such that

$$G(x_n - \rho_n \bar{x}) \leq G(x_n), \tag{15.3}$$

for all $n \geq n_0$.

Then $\min_{x \in X} G(x)$ admits a minimizer.

Proof: Let $\{\epsilon_n\} \subset (0,1)$ be a sequence converging to 0 from above and consider the family of auxiliary problems

$$\min_{x \in x} \; G(x) + \epsilon_n \, |x|^p. \tag{15.4}$$

By (i), every minimizing sequence for (15.1) is bounded. Extracting a weakly convergent subsequence, the existence of a solution $x_n \in X$ for (15.1) can be argued in a standard manner.

Suppose $\{x_n\}$ is bounded. Then there exists a weakly convergent subsequence, denoted by the same symbols and x^* such that $x_n \rightharpoonup x^*$. Passing to the limit $\epsilon_n \to 0^+$ in

$$G(x_n) + \frac{\epsilon_n}{2} |x_n|^p \leq G(x) + \frac{\epsilon_n}{2} |x|^p \text{ for all } x \in X$$

and using again (i) we have

$$G(x^*) \leq G(x) \text{ for all } x \in X$$

and thus x^* is a minimizer for G.

We now argue that $\{x_n\}$ is bounded, and assume to the contrary that $\lim_{n \to \infty} |x_n| = \infty$. Then a weakly convergent subsequence can be extracted from $\frac{x_n}{|x_n|}$ such that, again dropping indices, $\frac{x_n}{|x_n|} \rightharpoonup \bar{x}$, for some $\bar{x} \in X$. Since $G(x_n)$ is bounded from (ii)

$$\frac{x_n}{|x_n|} \to \bar{x} \text{ strongly in X.} \tag{15.5}$$

Next, choose $\rho > 0$ and n_0 according to (15.3) and thus for all $n \geq n_0$

$$G(x_n) + \epsilon_n \, |x_n|^p \leq G(x_n - \rho \bar{x}) + \epsilon_n \, |x_n - \rho \bar{x}|^p \leq G(x_n) + \epsilon_n \, |x_n - \rho \bar{x}|^p.$$

It follows that

$$|x_n| \leq |x_n - \rho \bar{x}| = |x_n - \rho_n \frac{x_n}{|x_n|} + \rho_n (\frac{x_n}{|x_n|} - \bar{x})| \leq |x_n| (1 - \frac{\rho}{|x_n|}) + \rho_n |\frac{x_n}{|x_n|} - \bar{x}|.$$

This implies that

$$1 \leq |\frac{x_n}{|x_n|} - \bar{x}|,$$

which give a contradiction to (15.5), and concludes the proof. □

The first half of condition (ii) is a compactness assumption and (15.3) is a compatibility condition [**?**].

<u>**Remark**</u> (1) If G is not bounded below. Define the (sequential) recession functional G_∞ associated with G. It is the extended real-valued functional defined by

$$G_\infty(x) = \inf_{x_n \rightharpoonup x, t_n \to \infty} \liminf_{n \to \infty} \frac{1}{t_n} G(t_n x_n). \tag{15.6}$$

and assume $G_\infty(x) \geq 0$. Then, $G(x) + \epsilon\,|x|^2$ is coercive. If not, there exists a sequence $\{x_n\}$ satisfying $G(x_n) + \epsilon\,|x_n|^2$ is bounded but $|x_n| \to \infty$. Let $w - \lim \frac{x_n}{|x_n|} = \bar{x}$.

$$\frac{1}{|x_n|} G(|x_n| \frac{x_n}{|x_n|}) + \epsilon |x_n| \to 0.$$

Since $G_\infty(\bar{x}) \leq \frac{1}{|x_n|} G(|x_n| \frac{x_n}{|x_n|})$, which yields a contradiction.

Moreover, in the proof note that

$$G_\infty(\bar{x}) \leq \liminf_{n\to\infty} \frac{1}{|x_n|} G(x_n) = \liminf_{n\to\infty} \frac{1}{|x_n|} G(\frac{x_n}{|x_n|}|x_n|) \leq 0.$$

If we assume $G_\infty \geq 0$, then $G_\infty(\bar{x}) = 0$. Thus, (15.3) must hold, assuming $G_\infty(\bar{x}) = 0$.

In [?] (15.3) is assumed for all $x \in X$ and $G_\infty(\bar{x}) = 0$, i.e.,

$$G(x - \rho\bar{x}) \leq G(x).$$

Thus, our condition is much weaker. In [?] we find a more general version of (15.3): there exists a ρ_n and z_n such that $\rho_n \in (0, |x_n|]$ and $z_n \in X$

$$G(x_n - \rho_n\, z_n) \leq G(x_n), \quad |z_n - z| \to 0 \text{ and } |z - \bar{x}| < 1.$$

(2) We have also proved that a sequence $\{x_n\}$ that minimizes the regularized problem for each ϵ_n has a weakly convergent subsequence to a minimizer of $G(x)$.

Example 1 If $ker(G_\infty) = 0$, then $\bar{x} = 0$ and the compactness assumption implies $|\bar{x}| = 1$, which is a contradiction. Thus, the theorem applies.

Example 2 (ℓ^0 optimization) Let $X = \ell^2$ and consider ℓ^0 minimization:

$$G(x) = \frac{1}{2}|Ax - b|_2^2 + \beta\,|x|_0, \tag{15.7}$$

where

$$|x|_0 = \text{the number of nonzero elements of } x \in \ell^2,$$

the counting measure of x. Assume $N(A)$ is finite and $R(A)$ is closed. The, one can apply Theorem to prove the existence of solutions to (15.7). That is, suppose $|x_n| \to \infty$, $G(x_n)$ is bounded from above and $\frac{x^n}{|x^n|} \rightharpoonup z$. First, we show that $z \in N(A)$ and $\frac{x^n}{|x^n|} \to z$. Since $\{G(x_n)\}$ is bounded from above, here exists M such that

$$J(x_n) = \frac{1}{2}|Ax_n - b|_2^2 + \beta\,|x_n|_0 \leq M, \text{ for all } n. \tag{15.8}$$

Since $R(A)$ closed, by the closed range theorem $X = R(A^*) + N(A)$ and let $x_n = x_n^1 + x_n^2$ Consequently, $0 \leq |Ax_n^1|_2 - 2(b, Ax_n^1)_2 + |b|_2^2 \leq K$ and $|Ax_n^1|$ is bounded.

$$0 \leq \Big|\, A(\frac{x_n^1}{|x_n|_2})\,\Big|_2^2 - 2\frac{1}{|x_n|_2}(\,A^*b, \frac{x_n}{|x_n|_2}\,)_2 + |b|_2^2 \to 0. \tag{15.9}$$

By the closed range theorem this implies that $|x_i^n$ is bounded and $\frac{x_n^1}{|x_n|_2} \to \bar{x}^1 = 0$ in ℓ^2. Since $\frac{x_n^2}{|x_n|_2} \rightharpoonup \bar{x}^2$ and by assumption $\dim N(A) < \infty$ it follows that $\frac{x_n}{|x_n|_2} \to z = \bar{x}^2$ strongly in ℓ^2. Next, (15.3) holds. Since

$$|x_n|_0 = |\frac{x_n}{|x_n|}|_0$$

and thus $|z|_0 < \infty$ (15.3) is equivalent to showing that

$$|(x_n^1 + x_n^2 - \rho\, z)_i|_0 \leq |(x_n^1 + x_n^2)_i|_0 \quad \text{for all } i,$$

and for all n sufficiently large. Only the coordinates for which $(x_n^1 + x_n^2)_i = 0$ with $z_i \neq 0$ require our attention. Since $|z|_0 < \infty$ there exists $\tilde{i}$ such that $z_i = 0$ for all $i > \tilde{i}$. For $i \in \{1, \ldots, \tilde{i}\}$ we define $\mathcal{I}_i = \{n : (x_n^1 + x_n^2)_i = 0,\ z_i \neq 0\}$. These sets are finite. In fact, if $\mathcal{I}_i$ is infinite for some $i \in \{1, \ldots, \tilde{i}\}$, then $\lim_{n\to\infty, n\in\mathcal{I}_i} \frac{1}{|x_n|_2}(x_n^1 + x_n^2)_i = 0$. Since $\lim_{n\to\infty} \frac{1}{|x_n|_2}(x_n^1)_i = 0$ this implies that $z_i = 0$, which is a contradiction.

Example 3 (Obstacle problem) Consider

$$\min \quad G(u) = \int_0^1 (\frac{1}{2}|u_x|^2 - fu)\, dx \tag{15.10}$$

subject to $u(\frac{1}{2}) \leq 1$. Let $X = H^1(0,1)$. If $|u_n| \to \infty$ and $v_n = \frac{u_n}{|u_n|} \rightharpoonup v$. By the assumption in (ii)

$$\frac{1}{2}\int_0^1 |(v_n)_x|^2\, dx \leq \frac{1}{|u_n|}\int_0^1 f v_n\, dx + \frac{G(u^n)}{|u_n|^2} \to 0$$

and thus $(v_n)_x \to 0$ and $v_n \to v = c$ for some constant c. But, since $v_n(\frac{1}{2}) \leq \frac{1}{|u_n|} \to 0$, $c = v(\frac{1}{2}) \leq 0$ Thus,

$$G(u_n - \rho\, v) = G(u_n) + \int_0^1 \rho v f(x)\, dx \leq G(u_n)$$

if $\int_0^1 f(x)\, dx \geq 0$. That is, if $\int_0^1 f(x)\, dx \geq 0$ (15.10) has a minimizer.

Example 4 (Friction Problem)

$$\min \quad \int_0^1 (\frac{1}{2}|u_x|^2 - fu)\, dx + |u(1)| \tag{15.11}$$

Using exactly the same argument above. we have $(v_n) \to 0$ and $v_n \to v = c$. But, we have

$$G(u_n - \rho v) = G(u_n) + \rho c \int_0^1 f(x)\, dx + |u_n(1) - \rho c| - |u_n(1)|$$

where

$$\frac{|u_n(1) - \rho c| - |u_n(1)|}{|u_n|} = |v_n(1) - \frac{\rho}{|u_n|}c| - |v_n(1)| \leq 0$$

for sufficiently large n since

$$\frac{|u_n(1) - \rho c| - |u_n(1)|}{|u_n|} \to |c - \hat{\rho} c| - |c| = -\hat{\rho}|c|$$

where $\hat{\rho} = \lim_{n\to\infty} \frac{\rho_n}{|u_n|}$. Thus, if $|\int_0^1 f(x)\, dx| \leq 1$, then (15.11) has a minimizer.

Example 5 (L^∞ Laplacian).

$$\min \quad \int_0^1 |u - f|\, dx \text{ subject to } |u_x| \leq 1$$

Similarly, we have $(v_n)_x \to 0$ and $v_n \to v = c$.

$$\frac{G(u_n - \rho c) - G(u_n)}{|u_n|} = G(v_n - \hat{\rho} c - \frac{f}{|u_n|}) - G(v_n - \frac{f}{|u_n|}) \leq 0$$

for sufficiently large n.

Example 6 (Elastic contact problem) Let $X = H^1(\Omega)^d$ and $\vec{u}$ is the deformation field. Given, the boundary body force $\vec{g}$, consider the elastic contact problem:

$$\min \quad \int_\Omega \frac{1}{2}\epsilon : \sigma \, dx - \int_{\Gamma_1} \vec{g} \cdot \vec{u} \, ds_x + \int_{\Gamma_2} |\tau \cdot \vec{u}| \, ds_x, \quad \text{subject to } n \cdot \vec{u} \leq \psi, \tag{15.12}$$

where we assume the linear strain:

$$\epsilon(u) = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})$$

and the Hooke's law:

$$\sigma = \mu\epsilon + \lambda \, tr\epsilon I$$

with positive Lame constants μ, λ. If $|\vec{u}_n| \to \infty$ and $G(u_n)$ is bounded we have

$$\epsilon(v_n) \to 0 \quad \nabla\epsilon(v_n) \to 0$$

for $\vec{v}_n = \frac{\vec{u}_n}{|\vec{u_n}|} \rightharpoonup v$. Thus, one can show that $\vec{v}_n \to v = (v^1, v^2) \in \{(-Ax_2 + C_1, Ax_1 + C_2)\}$ for the two dimensional problem. Consider the case when $\Omega = (0,1)^2$ and $\Gamma_1 = \{x_2 = 1\}$ and $\Gamma_2 = \{x_2 = 0\}$. As shown in the previous examples

$$\frac{G(\vec{u}_n - \rho \, v) - G(\vec{u}_n)}{|\vec{u}_n|} = \int_{\Gamma_1} \hat{\rho} n \cdot \vec{g} v^2 \, ds_x + \int_{\Gamma_1} \hat{\rho}\tau \cdot \vec{g} v^1 \, ds_x$$

$$+ \int_{\Gamma_2} (|(v_n)^1 - \hat{\rho} v^1| - |(v_n)^1|) \, ds_x$$

Since $v^2 \in C = \{Ax_1 + C_2 \leq 0\}$ and for $v^1 = C_1$

$$|(v_n)^1 - \hat{\rho} v^1| - |(v_n)^1| \to |(v)^1 - \hat{\rho} v^2| - |(v)^1| = -\hat{\rho}|v^1|,$$

if we assume $|\int_{\Gamma_1} \tau \cdot \vec{g} \, ds_x| \leq 1$ and $\int_{\Gamma_1} n \cdot \vec{g} v_2 \, ds_x \geq 0$ for all $v_2 \in C$, then (15.12) has a minimizer.

Example 7

$$\min \quad \int_0^1 (\frac{1}{2}(a(x)|u_x|^2 + |u|^2) - fu) \, dx$$

with $a(x) > 0$ except $a(\frac{1}{2}) = 0$.

15.2 Necessary Optimality

Let X be a Banach space and continuously embed to $H = L^2(\Omega)^m$. Let $J : X \to R^+$ is C^1 and N is nonsmooth functional on $H = L^2(\Omega)^m$ is given by

$$N(y) = \int_\Omega h(y(\omega)) \, d\omega.$$

where $h : R^m \to R$ is a lower semi-continuous functional. Consider the minimization on H;

$$\min \quad J(y) + N(y) \tag{15.13}$$

subject to

$$y \in \mathcal{C} = \{y(\omega) \in U \text{ a.e. }\}$$

where U is a closed convex set in R^m.

Given $y^* \in \mathcal{C}$, define needle perturbations of (the optimal solution) y^* by

$$y = \begin{cases} y^* & \text{on } |\omega - s| > \delta \\ u & \text{on } |\omega - s| \le \delta \end{cases}$$

for $\delta > 0$, $s \in \omega$, and $u \in U$. Assume that

$$|J(y) - J(y^*) - (J'(y^*), y - y^*)| \sim o(\text{meas}(\{|s| < \delta\}). \tag{15.14}$$

Let $\mathcal{H}$ is the Hamiltonian defined by

$$\mathcal{H}(y, \lambda) = (\lambda, y) - h(y), \quad y, \ \lambda \in R^m.$$

Then, we have the pointwise maximum principle

Theorem (Pointwise Optimality) If an integrable function y^* attains the minimum and (15.14) holds, then for a.e. $\omega \in \Omega$

$$\begin{aligned} &J'(y^*) + \lambda = 0 \\ &\mathcal{H}(u, \lambda(\omega)) \le \mathcal{H}(y^*(\omega), \lambda(\omega)) \quad \text{for all } u \in U. \end{aligned} \tag{15.15}$$

Proof: Since

$$\begin{aligned} &0 \ge J(y) + N(y) - (J(y^*) + N(y^*)) \\ &\qquad = J(y) - J(y^*) - (J'(y^*), y - y^*) + \int_{|\omega - s| < \delta} (J'(y^*), u - y^*) + h(u) - h(y^*))\, d\omega \end{aligned}$$

Since h is lower semicontinuos, it follows from (15.14) that at a Lubegue point $s = \omega$ of y^* (15.15) holds. □

For $\lambda \in R^m$ define a set-valued function by

$$\Phi(\lambda) = \text{argmax}_{y \in U}\{(\lambda, y) - h(y)\}.$$

Then, (15.15) is written as

$$y(\omega) \in \Phi(-J'(y(\omega))$$

The graph of Φ is monotone:

$$(\lambda_1 - \lambda_2, y_1 - y_2) \ge 0 \text{ for all } y_1 \in \Phi(\lambda_1), \ \ y_2 \in \Phi(\lambda_2).$$

but it not necessarily maximum monotone. Let h^* be the conjugate function of h defined by

$$h^*(\lambda) = \max_{u \ inU} \mathcal{H}(\lambda, y).$$

Then, h^* is necessarily convex. Since if $u \in \Phi(\lambda)$, then

$$h^*(\hat{\lambda}) \geq (\hat{\lambda}, y) - h(u) = (\hat{\lambda} - \lambda, y) + h^*(u),$$

thus $u \in \partial h^*(\lambda)$. Since h^* is convex, ∂h^* is maximum monotone and thus ∂h^* is the maximum monotone extension.

Let h^{**} be the bi-conjugate function of h;

$$h^{**}(x) = \sup_\lambda \{(x, \lambda) - \sup_y \mathcal{H}(y, \lambda)\}$$

whose epigraph is the convex envelope of the epigraph of h, i.e.,

$$epi(h^{**}) = \{\lambda_1 \, y_1 + (1 - \lambda) \, y_2 \text{ for all } y_1, \; y_2 \in epi(h) \text{ and } \lambda \in [0, 1]\}.$$

Then, h^{**} is necessarily convex and is the convexfication of h and

$$\lambda \in \partial h^{**}(u) \text{ if and only if } u \in \partial h^*(\lambda)$$

and

$$h^*(\lambda) + h^{**}(u) = (\lambda, u).$$

Consider the relaxed problem:

$$\min \quad J(y) + \int_\Omega h^{**}(y(\omega)) \, d\omega. \tag{15.16}$$

Since h^{**} is convex,

$$N^{**}(y) = \int_\Omega h^{**}(y) \, d\omega$$

is weakly lower semi-continuous and thus there exits a minimizer y of (15.16). It thus follows from Theorem (Pointwise Optimality) that the necessary optimality of (15.16) is given by

$$-J'(y)(\omega) \in \partial h^{**}(y(\omega)),$$

or equivalently

$$y \in \partial h^*(-J'(y)). \tag{15.17}$$

Assume

$$J(\hat{y}) - J(y) - J'(y)(\hat{y} - y) \geq 0 \text{ for all } \hat{y} \in \mathcal{C},$$

where

$$y(\omega) \in \Phi(-J'(y(\omega))).$$

Then, y minimizes the original cost functional (15.13).

One can construct a C^1 realization for h by

$$h^*_\epsilon(\lambda) = \min_p \{\frac{1}{2\epsilon}|p - \lambda|^2 + h^*(p)\}.$$

That is, h^*_ϵ is C^1 and $(h^*_\epsilon)'$ is Lipschitz and $(h^*_\epsilon)'$ is the Yosida approximation of ∂h^*. Then, we let $h_\epsilon = (h^*_\epsilon)^*$

$$h_\epsilon(u) = \min\{\frac{1}{2\epsilon}|y - u|^2 + h(y)\}.$$

Consider the relaxed problem

$$\min \quad J(y) + \int_\Omega h_\epsilon(y(\omega))\,d\omega. \tag{15.18}$$

The necessary optimality condition becomes

$$y(\omega) = (h_\epsilon^*)'(-J'(y(\omega))). \tag{15.19}$$

Then, one can develop the semi smooth Newton method to solve (15.19).

Since

$$|h_\epsilon - h^{**}|_\infty \le C h^{**}\,\epsilon,$$

one can argue that a sequence of minimizers y_ϵ of (15.18) to a minimizer of (15.16). In fact,

$$J(y_\epsilon) + \int_\Omega h_\epsilon(y_\epsilon)\,d\omega. \le J(y) + \int_\Omega h_\epsilon(y)\,d\omega.$$

for all $y \in \mathcal{C}$ Suppose $|y_\epsilon|$ is bounded , then y_ϵ has a weakly convergent subsequence to $\bar{y} \in \mathcal{C}$ and

$$J(\bar{y}) + \int_\Omega h^{**}(\bar{y})\,d\omega. \le J(y) + \int_\Omega h^{**}(y)\,d\omega.$$

for all $y \in \mathcal{C}$.

Example ($L^0(\Omega)$ optimization) Based on our analysis one can analyze when h involves the volume control by $L^0(\Omega)$, i.e., Let h on R be given by

$$h(u) = \frac{\alpha}{2}|u|^2 + \beta|u|_0, \tag{15.20}$$

where

$$|u|_0 = \begin{cases} 0 \text{ if } u = 0 \\ 1 \text{ if } u \neq 0. \end{cases}$$

That is,

$$\int_\Omega |u|_0\,d\omega = \text{meas}(\{u(\omega) \neq 0\}).$$

In this case Theorem applies with

$$\Phi(q) := argmin_{u\in R}(h(u) - qu) = \begin{cases} \frac{q}{\alpha} & \text{for } |q| \ge \sqrt{2\alpha\beta} \\ 0 & \text{for } |q| < \sqrt{2\alpha\beta}. \end{cases} \tag{15.21}$$

and the conjugate function h^* of h;

$$-h^*(q) = h(\Phi(q)) - q\,\Phi(q) = \begin{cases} -\frac{1}{2\alpha}|q|^2 + \beta & \text{for } |q| \ge \sqrt{2\alpha\beta} \\ 0 & \text{for } |q| < \sqrt{2\alpha\beta}. \end{cases}$$

The bi-conjugate function $h^{**}(u)$ is given by

$$h^{**}(u) = \begin{cases} \frac{\alpha}{2}|u|^2 + \beta & |u| > \sqrt{\frac{2\beta}{\alpha}} \\ \sqrt{2\alpha\beta}\,|u| & |u| \le \sqrt{\frac{2\beta}{\alpha}}. \end{cases}$$

Clearly, $\Phi : R \to R$ is monotone, but it is not maximal monotone. The maximal monotone extension $\tilde{\Phi} = (\partial h^{**})^{-1} = \partial h^*$ of Φ and is given by

$$\tilde{\Phi}(q) \in \begin{cases} \frac{q}{\alpha} & \text{for } |q| > \sqrt{2\alpha\beta} \\ \\ 0 & \text{for } |q| < \sqrt{2\alpha\beta} \\ \\ [\frac{q}{\alpha}, 0] & \text{for } q = -\sqrt{2\alpha\beta} \\ \\ [0, \frac{q}{\alpha}] & \text{for } q = \sqrt{2\alpha\beta}. \end{cases}$$

Thus,

$$(h^*_\epsilon)' = \frac{\lambda - p_\epsilon(\lambda)}{\epsilon}$$

where

$$p_\epsilon(\lambda) = \text{argmin}\{\frac{1}{2\epsilon}|p-\lambda|^2 + h^*(p)\}.$$

We have

$$p_\epsilon(\lambda) = \begin{cases} \lambda & |\lambda| < \sqrt{2\alpha\beta} \\ \\ \sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} \le \lambda \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \\ -\sqrt{2\alpha\beta} & \sqrt{2\alpha\beta} \le -\lambda \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \\ \frac{\lambda}{1+\frac{\epsilon}{\alpha}} & |\lambda| \ge (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}. \end{cases} \tag{15.22}$$

$$(h^*_\epsilon)'(\lambda) = \begin{cases} 0 & |\lambda| < \sqrt{2\alpha\beta} \\ \\ \frac{\lambda-\sqrt{2\alpha\beta}}{\epsilon} & \sqrt{2\alpha\beta} \le \lambda \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \\ \frac{\lambda+\sqrt{2\alpha\beta}}{\epsilon} & \sqrt{2\alpha\beta} \le -\lambda \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \\ \frac{\lambda}{\alpha+\epsilon} & |\lambda| \ge (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}. \end{cases}$$

Example (Authority optimization) We analyze the authority optimization of two unknowns u_1 and u_2, i.e., the at least one of u_1 and u_2 is zero. We formulate such a problem by using

$$h(u) = \frac{\alpha}{2}(|u|_1^2 + |u_2|^2) + \beta|u_1u_2|_0, \tag{15.23}$$

The, we have

$$h^*(p_1, p_2) = \begin{cases} \frac{1}{2\alpha}(|p_1|^2 + |p_2|^2) - \beta & |p_1|,\ |p_2| \ge \sqrt{2\alpha\beta} \\ \\ \frac{1}{2\alpha}|p_1|^2 & |p_1| \ge |p_2|,\ \ |p_2| \le \sqrt{2\alpha\beta} \\ \\ \frac{1}{2\alpha}|p_2|^2 & |p_2| \ge |p_1|,\ \ |p_1| \le \sqrt{2\alpha\beta} \end{cases}$$

Also,we have

$$p_\epsilon(\lambda_1,\lambda_2)=\begin{cases}\frac{1}{1+\frac{\epsilon}{\alpha}}(\lambda_1,\lambda_2) & |\lambda_1|,\ |\lambda_2|\ge(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}\\ (\frac{1}{1+\frac{\epsilon}{\alpha}}\lambda_1,\pm\sqrt{2\alpha\beta}) & |\lambda_1|\ge(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta}\le|\lambda_2|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}\\ (\pm\sqrt{2\alpha\beta},\frac{1}{1+\frac{\epsilon}{\alpha}}\lambda_2) & |\lambda_2|\ge(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta}\le|\lambda_1|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}\\ (\lambda_2,\lambda_2) & |\lambda_2|\le|\lambda_1|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta}\le|\lambda_2|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}\\ (\lambda_1,\lambda_1) & |\lambda_1|\le|\lambda_2|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta}\le|\lambda_1|\le(1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta}\\ (\frac{\lambda_1}{1+\frac{\epsilon}{\alpha}},\lambda_2) & |\lambda_1|\ge(1+\frac{\epsilon}{\alpha})|\ge|\lambda_2|,\ \ |\lambda_2|\le\sqrt{2\alpha\beta}\\ (\lambda_1,\frac{\lambda_2}{1+\frac{\epsilon}{\alpha}}) & |\lambda_2|\ge(1+\frac{\epsilon}{\alpha})|\ge|\lambda_1|,\ \ |\lambda_1|\le\sqrt{2\alpha\beta}\\ (\lambda_2,\lambda_2) & |\lambda_2|\le|\lambda_1|\le(1+\frac{\epsilon}{\alpha})|\lambda_2|,\ \ |\lambda_2|\le\sqrt{2\alpha\beta}\\ (\lambda_1,\lambda_1) & |\lambda_1|\le|\lambda_2|\le(1+\frac{\epsilon}{\alpha})|\lambda_1|,\ \ |\lambda_1|\le\sqrt{2\alpha\beta}\end{cases}\tag{15.24}$$

15.3 Augmented Lagrangian method

Given $\lambda\in H$ and $c>0$, consider the augmented Lagrangian functional;

$$L(x,y,\lambda)=J(x)+(\lambda,x-y)_H+\frac{c}{2}|x-y|_H^2+N(y).$$

Since

$$L(x,y,\lambda)=J(x)+\int_\Omega(\frac{c}{2}|x-y|^2+(\lambda,x-y)+h(y))\,d\omega$$

it follows from Theorem that if $(\bar{x},\bar{y})$ minimizes $L(x,y,\lambda)$, then the necessary optimality condition is given by

$$-J'(\bar{x})=\lambda+c\,(\bar{x}-\bar{y})$$

$$\bar{y}\in\Phi_c(\lambda+c\,\bar{x}).$$

where

$$\Phi_c(p)=\text{argmax}\{(p,y)-\frac{c}{2}|y|^2-h(y)\}$$

Thus, the augmented Lagrangian update is given by

$$\lambda^{n+1}=\lambda^n+c\,(x^n-y^n)$$

where (x^n,y^n) solves

$$-J'(x^n)=\lambda^n+c\,(x^n-y^n)$$

$$y^n\in\Phi_c(\lambda^n+c\,x^n).$$

If we convexity N by the bi-conjugate of h we have

$$L(x,y,\lambda)=J(x)+(\lambda,x-y)_H+\frac{c}{2}|y-x|_H^2+N^{**}(y).$$

Since

$$h_c(x,\lambda)=\inf_y\{h^{**}(y)+(\lambda,x-y)+\frac{c}{2}\,|y-x|^2\}$$

$$=\inf_y\{\sup_p\{(y,p)-h^*(p)\}+(\lambda,x-y)+\frac{c}{2}\,|y-x|^2\}$$

$$=\sup_p\{-\frac{1}{2c}|p-\lambda|^2+(x,p)-h^*(p)\}.$$

we have

$$h_c'(x,\lambda)=\text{argmax}\{-\frac{1}{2c}|p-\lambda|^2+(x,p)-h^*(p)\}.$$

If we let

$$p_c(\lambda)=\text{argmin}\{\frac{1}{2c}|p-\lambda|^2+h^*(p)\},$$

then

$$h_c'(x,\lambda)=p_c(\lambda+cx) \tag{15.25}$$

and

$$(h_c^*)'(\lambda)=\frac{\lambda-p_c(\lambda)}{c}. \tag{15.26}$$

Thus, the maximal monotone extension $\tilde{\Phi}_c$ of the graph Φ_c is given by $(h_c^*)'(\lambda)$, the Yosida approximation of $\partial h^*(\lambda)$ and results in the update:

$$-J'(x^n)=\lambda^n+c\,(x^n-y^n)$$

$$y^n=(h_c^*)'(\lambda^n+c\,x^n).$$

In the case $c=0$, we have the multiplier method

$$\lambda^{n+1}=\lambda^n+\alpha\,(x^n-y^n),\ \alpha>0$$

where (x^n,y^n) solves

$$-J'(x^n)=\lambda^n$$

$$y^n\in\tilde{\Phi}(\lambda^n).$$

In the case $\lambda=0$ we have the proximal iterate of the form

$$x^{n+1}=(h_c^*)'(x^n+\alpha\,J'(x^n)).$$

15.4 Semismooth Newton method

The necessary optimality condition is written as

$$-J'(x)=\lambda$$

$$\lambda=h_c'(x,\lambda)=p_c(\lambda+c\,x).$$

is the semi-smooth equation. For the case of $h = \frac{\alpha}{2}|x|^2 + \beta|x|_0$ we have from (15.22)-(15.25)

$$h_c'(x,\lambda) = \begin{cases} \lambda & \text{if } |\lambda + cx| \le \sqrt{2\alpha\beta} \\ \sqrt{2\alpha\beta} & \text{if } \sqrt{2\alpha\beta} \le \lambda + cx \le (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ -\sqrt{2\alpha\beta} & \text{if } \sqrt{2\alpha\beta} \le -(\lambda + cx) \le (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ \frac{\lambda}{\alpha+c} & \text{if } |\lambda + cx| \ge (1+\frac{c}{\alpha})\sqrt{2\alpha\beta}. \end{cases}$$

Thus, the semi-smooth Newton update is

$$-J'(x^{n+1}) = \lambda^{n+1}, \quad \begin{cases} x^{n+1} = 0 & \text{if } |\lambda^n + cx^n| \le \sqrt{2\alpha\beta} \\ \lambda^{n+1} = \sqrt{2\alpha\beta} & \text{if } \sqrt{2\alpha\beta} < \lambda^n + cx^n < (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ \lambda^{n+1} = -\sqrt{2\alpha\beta} & \text{if } \sqrt{2\alpha\beta} < -(\lambda^n + cx^n) < (1+\frac{c}{\alpha})\sqrt{2\alpha\beta} \\ x^{n+1} = \frac{\lambda^{n+1}}{\alpha} & \text{if } |\lambda^n + cx^n| \ge (1+\frac{c}{\alpha})\sqrt{2\alpha\beta}. \end{cases}$$

For the case $h(u_1,u_2) = \frac{\alpha}{2}(|u_1|^2 + |u_2|^2) + \beta|u_1u_2|_0$. Let $\mu^n = \lambda^n + cu^n$. From (15.24) the semi-smooth Newton update is

$$\begin{cases} u^{n+1} = \frac{\lambda^{n+1}}{\alpha} & |\mu_1^n|,\ |\mu_2^n| \ge (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ u_1^{n+1} = \frac{\lambda_1^{n+1}}{\alpha},\ \ \lambda_2^{n+1} = \pm\sqrt{2\alpha\beta} & |\mu_1^n| \ge (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta} \le |\mu_2^n| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \lambda^{n+1} = \pm\sqrt{2\alpha\beta},\ \ u_2^{n+1} = \frac{\lambda_2^{n+1}}{\alpha} & |\mu_2^n| \ge (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta} \le |\mu_1^n| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \lambda_1^{n+1}, \lambda_2^{n+1},\ \ u_1^{n+1} = \frac{\lambda_1^{n+1}}{\alpha} & |\mu_2^n| \le |\mu_1^n| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta} \le |\mu_2^n| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ \lambda_1^{n+1} = \lambda_2^{n+1},\ \ u_2^{n+1} = \frac{\lambda_2^{n+1}}{\alpha} & |\lambda_1| \le |\lambda_2| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta},\ \ \sqrt{2\alpha\beta} \le |\lambda_1| \le (1+\frac{\epsilon}{\alpha})\sqrt{2\alpha\beta} \\ u_1^{n+1} = \frac{\lambda_1^{n+1}}{\alpha},\ \ u_2^{n+1} = 0 & |\mu_1^n| \ge (1+\frac{\epsilon}{\alpha})| \ge |\mu_2^n|,\ \ |\mu_2^n| \le \sqrt{2\alpha\beta} \\ u_2^{n+1} = \frac{\lambda_2^{n+1}}{\alpha},\ \ u_1^{n+1} = 0 & |\mu_2^n| \ge (1+\frac{\epsilon}{\alpha})| \ge |\mu_1^n|,\ \ |\mu_1^n| \le \sqrt{2\alpha\beta} \\ \lambda_2^{n+1} = \lambda_2^{n+1}\ \ u_2^{n+1} = 0 & |\mu_2^n| \le |\mu_1^n| \le (1+\frac{\epsilon}{\alpha})|\mu_2^n|,\ \ |\mu_2^n| \le \sqrt{2\alpha\beta} \\ \lambda_1^{n+1} = \lambda_2^{n+1},\ \ u_1^{n+1} = 0 & |\mu_1^n| \le |\mu_2^n| \le (1+\frac{\epsilon}{\alpha})|\mu_1^n|,\ \ |\mu_1^n| \le \sqrt{2\alpha\beta}. \end{cases}$$

15.5 Constrained optimization

Consider the constrained minimization problem of the form;

$$\min \quad J(x,u) = \int_\Omega (\ell(x) + h(u))\, d\omega \text{ over } u \in \mathcal{U}, \tag{15.27}$$

subject to the equality constraint

$$\tilde{E}(x,u) = Ex + f(x) + Bu = 0. \tag{15.28}$$

Here $\mathcal{U} = \{u \in L^2(\Omega)^m : u(\omega) \in U \text{ a.e. in } \Omega\}$, where U is a closed convex subset of R^m. Let X is a closed subspace of $L^2(\Omega)^n$ and $E : X \times \mathcal{U} \to X^*$ with $E \in \mathcal{L}(X, X^*)$ bounded invertible, $B \in \mathcal{L}(L^2(\Omega)^m, L^2(\Omega)^n)$ and $f : L^2(\Omega)^n \to L^2(\Omega)^n$ Lipschitz. We assume $\tilde{E}(x,u) = 0$ has a unique solution $x = x(u) \in X$, given $u \in \mathcal{U}$. We assume there exist a solution $(\bar{x}, \bar{x})$ to Problem (15.32). Also, we assume that the adjoint equation:

$$(E + f_x(\bar{x}))^* p + \ell'(\bar{x}) = 0. \tag{15.29}$$

has a solution in X. It is a special case of (15.1) when $J(u)$ is the implicit functional of u:

$$J(u) = \int_\Omega \ell(x(\omega))\, d\omega,$$

where $x = x(u) \in X$ for $u \in \mathcal{U}$ is the unique solution to $Ex + f(x) + Bu = 0$.

To derive a necessary condition for this class of (nonconvex) problems we use a maximum principle approach. For arbitrary $s \in \Omega$, we shall utilize needle perturbations of the optimal solution $\bar{u}$ defined by

$$v(\omega) = \begin{cases} u & \text{on } \{\omega : |\omega - s| < \delta\} \\ \\ \bar{u}(\omega) & \text{otherwise,} \end{cases} \tag{15.30}$$

where $u \in U$ is constant and $\delta > 0$ is sufficiently small so that $\{|\omega - s| < \delta\} \subset \Omega$. The following additional properties for the optimal state $\bar{x}$ and each perturbed state $x(v)$ will be used:

$$\begin{cases} |x(v) - \bar{x}|^2_{L^2(\Omega)} = o(\text{meas}(\{|\omega - s| < \delta\}))^{\frac{1}{2}} \\ \\ \int_\Omega \big(\, \ell(\cdot, x(v)) - \ell(\cdot, \bar{x}) - \ell_x(\cdot, \bar{x})(x(v) - \bar{x})\big) d\omega \, = O(|x(v) - \bar{x}|^2_{L^2(\Omega)}) \\ \\ \langle f(\cdot, x(v)) - f(\cdot, \bar{x}) - f_x(\cdot, \bar{x})(x(v) - \bar{x}), p\rangle_{X^*, X} = O(|x(v) - \bar{x}|^2_X) \end{cases} \tag{15.31}$$

Theorem Suppose $(\bar{x}, \bar{u}) \in X \times \mathcal{U}$ is optimal for problem (15.32) , that $p \in X$ satisfies the adjoint equation (15.29) and that (**??**), (**??**), and (15.31) hold. Then we have the necessary optimality condition that $\bar{u}(\omega)$ minimizes $h(u) + (p, Bu)$ pointwise a.e. in Ω.
Proof: By the second property in (15.31) we have

$$0 \le J(v) - J(\bar{u}) = \int_\Omega \big(\ell(\cdot, x(v)) - \ell(\cdot, x(\bar{u})) + h(v) - h(\bar{u})\big)\, d\omega$$

$$= \int_\Omega (\ell_x(\cdot, \bar{x})(x - \bar{x}) + h(v) - h(\bar{u}))\ d\omega + O(|x - \bar{x}|^2),$$

where

$$(\ell_x(\cdot, \bar{x})(x - \bar{x}), p) = -\langle (E + f_x(\cdot, \bar{x})(x - \bar{x})), p\rangle$$

and

$$0 = \langle E(x - \bar{x}) + f(\cdot, x) - f(\cdot, \bar{x}) + B(v - \bar{u}), p\rangle$$

$$= \langle E(x - \bar{x}) + f_x(\cdot, \bar{x})(x - \bar{x}) + B(v - \bar{u}), p\rangle + O(|x - \bar{x}|^2)$$

Thus, we obtain

$$0 \le J(v) - J(\bar{u}) \le \int_\Omega ((-B^*p, v - \bar{u}) + h(v) - h(\bar{u}))\, d\omega.$$

Dividing this by $meas(\{|\omega - s| < \delta\}) > 0$, letting $\delta \to 0$, and using the first property in (15.31) we obtain the desired result at a Lebesgue point s

$$\omega \to (-B^*p, v - \bar{u}) + h(v) - h(\bar{u}))$$

since h is lower semi continuous. □

Consider the relaxed problem of (15.32):

$$\min \quad J(x,u) = \int_\Omega (\ell(x) + h^{**}(u))\, d\omega \text{ over } u \in \mathcal{U}, \tag{15.32}$$

We obtain the pointwise necessary optimality

$$\begin{cases} E\bar{x} + f'(\bar{x}) + B\bar{u} = 0 \\ \\ (E + f'(\bar{x}))^* p + \ell'(\bar{x}) = 0 \\ \\ \bar{u} \in \partial h^*(-B^*p), \end{cases}$$

or equivalently

$$\lambda = h_c'(u, \lambda), \quad \lambda = B^*p.$$

where h_c is the Yoshida-M approximation of h^{**}.

15.6 Algoritm

Based the complementarity condition (**??**) we have the Primal dual active method for L^p optimization ($h(u) = \frac{\alpha}{2}|u|^2 + \beta\,|u|^p.$):

Primal-Dual Active method ($L^p(\Omega)$-optimization)

1. Initialize $u^0 \in X$ and $\lambda^0 \in H$. Set $n = 0$.
2. Solve for $(y^{n+1}, u^{n+1}, p^{n+1})$

$$Ey^{n+1} + Bu^{n+1} = g, \;\; E^*p^{n+1} + \ell'(y^{n+1}) = 0, \;\; \lambda^{n+1} = B^*p^{n+1}$$

   and

$$u^{n+1} = -\frac{\lambda^{n+1}}{\alpha + \beta p |u^n|^{p-2}} \text{ if } \omega \in \{|\lambda^n + c\,u^n| > (\mu_p + cu_p\}$$

$$\lambda^{n+1} = \mu_p \text{ if } \omega \in \{\mu_p \le \lambda^n + cu^n \le \mu_p + cu_n\}.$$

$$\lambda^{n+1} = -\mu_p \text{ if } \omega \in \{\mu_p \le -(\lambda^n + cu^n) \le \mu_p + cu_n\}.$$

$$u^{n+1} = 0, \text{ if } \omega \in \{|\lambda^n + cu^n| \le \mu_p\}.$$

3. Stop, or set $n = n + 1$ and return to the second seep.

16 Exercise

Problem 1 If A is a symmetric matrix on R^n and $b \in R^n$ is a vector. Let $F(x) = \frac{1}{2}x^tAx - b^tx$. Show that $F'(x) = Ax - b$.

Problem 2 Sequences $\{f_n\}$, $\{g_n\}$, $\{h_n\}$ are Cauchy in $L^2(0,1)$ but not on $C[0,1]$.
Problem 3 Show that $||x| - |y|| \le |x - y|$ and thus the norm is continuous.
Problem 4 Show that all norms are equivalent on a finite dimensional vector space.

Problem 5 Show that for a closed linear operator T, the domain $\mathcal{D}(T)$ is a Banach space if it is equipped by the graph norm

$$|x|_{\mathcal{D}(T)} = (|x|_X^2 + |Tx|_Y^2)^{1/2}.$$

Problem 6 Consider the spline interpolation problem (3.8). Find the Riesz representations F_i and $\tilde{F}_j$ in $H_0^2(0,1)$. Complete the spline interpolation problem (3.8).

Problem 7 Derive the necessary optimality for L^1 optimization (3.32) with $U = [-1,1]$.

Problem 8 Solve the constrained minimization in a Hilbert space X;

$$\min \quad \frac{1}{2}|x|^2 - (a,x)$$

subject to

$$(y_i, x) = b_i,\ 1 \le i \le m_1, \quad (\tilde{y}_j, x) \le c_j,\ 1 \le j \le m_2.$$

over $x \in X$. We assume $\{y_i\}_{i=1}^{m_1}$, $\{\tilde{y}_j\}_{j=1}^{m_2}$ are linearly independent.

Problem 9 Solve the pointwise obstacle problem:

$$\min \quad \int_0^1 (\frac{1}{2}|\frac{du}{dx}|^2 - f(x)u(x))\,dx$$

subject to

$$u(\frac{1}{3}) \le c_1, \quad u(\frac{2}{3}) \le c_2,$$

over a Hilbert space $X = H_0^1(0,1)$ for $f = 1$.

Problem 10 Find the differential form of the necessary optimality for

$$\min \quad \int_0^1 (\frac{1}{2}(a(x)|\frac{du}{dx}|^2 + c(x)|u|^2) - f(x)u(x))\,dx + \frac{1}{2}(\alpha_1|u(0)|^2 + \alpha_2|u(1)|^2) - (g_1, u(0)) - (g_2, u(1))$$

over a Hilbert space $X = H^1(0,1)$.

Problem 11 Let $B \in R^{m,n}$. Show that $R(B) = R^m$ if and only if BB^t is positive definite on R^m. Show that $x^* = B^t(BB^t)^{-1}c$ defines a minimum norm solution to $Bx = c$.

Problem 12 Consider the optimal control problem (3.44) (Example Control problem) with $\hat{U} = \{u \in R^m : |u|_\infty \le 1\}$. Derive the optimality condition (3.45).

Problem 13 Consider the optimal control problem (3.44) (Example Control problem) with $\hat{U} = \{u \in R^m : |u|_2 \le 1\}$. Find the orthogonal projection $P_\mathcal{C}$ and derive the necessary optimality and develop the gradient method.

Problem 14 Let $J(x) = \int_0^1 |x(t)|\,dt$ is a functional on $C(0,1)$. Show that $J'(d) = \int_0^1 \text{sign}(x(t))d(t)\,dt$.

Problem 15 Develop the Newton methods for the optimal control problem (3.44) without the constraint on u (i.e., $\hat{U} = R^m$).

Problem 16 (1) Show that the conjugate functional

$$h^*(p) = \sup_u \{(p,u) - h(u)\}$$

is convex.
(2) The set-valued function Φ

$$\Phi(p) = \text{argmax}_u \{(p,u) - h(u)\}$$

is monotone.

Problem 17 Find the graph $\Phi(p)$ for

$$\max_{|u| \le \gamma} \{pu - |u|_0\}$$

and

$$\max_{|u| \le \gamma} \{pu - |u|_1\}.$$