December 7, 2018 · · · Sorbonne Université, France

Finite codimensionality approach for general optimization problems and applications

Xu Zhang

School of Mathematics
Sichuan University, China

zhang_xu@scu.edu.cn

(Based on an ongoing joint work with Xu Liu, Qi Lü and Haisen Zhang)

Outline

1. Review on elementary facts
2. General optimization problems
3. Characterization of finite codimensionality
4. Application to deterministic optimal control problems
5. Application to stochastic optimal control problems

1. Review on elementary facts

$\diamondsuit$ Free extremum: Consider a minimizer x_0 of a smooth function $f^0(\cdot)$ in a set $G \subset \mathbb{R}^n$ (for some $n \in \mathbb{N}$), i.e., x_0 satisfies

$$f^0(x_0) = \inf_{x \in G} f^0(x). \tag{1}$$

If a nonzero vector $\ell \in \mathbb{R}^n$ is admissible for x_0 (i.e., there is a $\delta > 0$ so that $x_0 + s\ell \in G$ for any $s \in [0, \delta]$), then one has the following first-order necessary condition:

$$0 \le \lim_{s \to 0^+} \frac{f^0(x_0 + s\ell) - f^0(x_0)}{s} = \langle f_x^0(x_0), \ell \rangle. \tag{2}$$

In particular,

(i) If G is open, then

$$f_x^0(x_0) = 0.$$

(ii) If G is convex, one has

$$0 \le \langle f_x^0(x_0), x - x_0 \rangle, \quad \forall\, x \in G. \tag{3}$$

$\diamondsuit$ Extremum with constraints: Suppose $f : G \to \mathbb{R}^m$ (for some $m \in \mathbb{N}$) is another given function. Consider a minimizer x_0 of $f^0(\cdot)$ in G under the constraint $f(\cdot) = 0$, i.e., x_0 satisfies

$$f^0(x_0) = \inf_{x \in G} \{ f^0(x) \mid f(x) = 0 \}. \tag{4}$$

Introduce the following Lagrange function:

$$L(x) = f^0(x) + \langle \lambda, f(x) \rangle,$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multiplier.

If a nonzero vector $\ell \in \mathbb{R}^n$ is admissible for x_0, then one has the following first-order necessary condition:

$$\begin{cases} 0 \leq \langle f^0_x(x_0) + \lambda^\top f_x(x_0), \ell \rangle, \\ f(x_0) = 0. \end{cases} \tag{5}$$

◊ What should be done next?

• More general setting, especially the problems in infinite dimensions.

• Usually, in the general setting, the Implicit Function Theorem does not work well and therefore the constraints cannot be essentially removed. People then introduce Generalized Lagrange Multipliers.

• In order to guarantee the nontriviality of Generalized Lagrange Multipliers, people further introduce the so-called finite codimensionality condition.

• It is very difficult to verify the above condition directly except for some very special cases.

• A key contribution in this work is to reduce further the above mentioned condition to a suitable *a priori* estimate.

2. General optimization problems

Let V be a reflexive Banach space, X be a Hilbert space and E be a closed subset of X. Assume that two functions $f : \mathcal{D}(f) \subseteq V \to X$ and $f^0 : \mathcal{D}(f^0) \subseteq V \to \mathbb{R}$ satisfy suitable assumptions.

Consider the following general optimization problem:

$$(\mathbf{P}) \qquad \inf_{u \in \mathcal{D}(f) \cap \mathcal{D}(f^0)} \left\{ f^0(u) \;\middle|\; f(u) \in E \right\}.$$

Assume that $\overline{u} \in \mathcal{D}(f) \cap \mathcal{D}(f^0)$ is a solution to the optimization problem $(\mathbf{P})$ and set $\overline{e} = f(\overline{u})$.

In the following, some first-order necessary condition on $\overline{u}$ will be given.

$\diamondsuit$ Some notions

• A set $I_E(e) \subseteq X$ is called the intermediate cone to E at $e \in E$, if for any $w \in I_E(e)$ and $\{h_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ with $h_k \searrow 0$ as $k \to \infty$, $\exists$ $\{e_k\}_{k=1}^{\infty} \subseteq E$, such that

$$\lim_{k\to\infty} e_k = e \quad \text{and} \quad \lim_{k\to\infty} \frac{e_k - e}{h_k} = w.$$

• A vector $\xi \in X$ is called a variation of $g : \mathcal{D}(g) \subseteq V \to X$ at $e \in \mathcal{D}(g)$, if $\exists$ $\{h_k\}_{k=1}^{\infty} \subseteq \mathbb{R}$ with $h_k \searrow 0$ as $k \to \infty$, and $\{e_k\}_{k=1}^{\infty} \subseteq \mathcal{D}(g)$, such that

$$|e_k - e| \leq h_k \quad \text{and} \quad \lim_{k\to\infty} \frac{g(e_k) - g(e)}{h_k} = \xi.$$

The set of all variations of g at e is denoted by Var $g(e)$.

• If $g : \mathbb{R} \to \mathbb{R}$ is differentiable at $e \in \mathbb{R}$, then

$$\text{Var } g(e) = \left\{ \xi \in \mathbb{R} \;\middle|\; \xi = g'(e)x \text{ for some } x \in [-1, 1] \right\}.$$

• A linear subspace X_0 of X is called finite codimensional, if $\exists\, m \in \mathbb{N}$ and linearly independent $x_1, x_2, \cdots, x_m \in X \setminus X_0$, such that

$$\text{span}\{X_0, x_1, x_2, \cdots, x_m\} = X.$$

• A subset D of X is called finite codimensional in X, if

$(\mathbf{1})$ $\exists\, x_0 \in \overline{\text{co}}\, D$, such that span $\{D - x_0\}$ is a finite codimensional subspace of X; and

$(\mathbf{2})$ $\overline{\text{co}}\,(D - y_0)$ has at least one interior point in this subspace.

♢ First-order necessary condition

Theorem 1: *Assume that for some* $\rho > 0$,

(H) *Var* $f(\overline{u}) - I_E(\overline{e}) \cap B_\rho(0)$ *is finite codimensional in* X.

Then $\exists$ *nonzero* $(z_0, z) \in \mathbb{R}^+ \times X$ *such that*

$$\begin{cases} z_0\xi_0 + (z, \xi)_X \geq 0, \qquad \forall\, (\xi_0, \xi) \in \textit{Var}\,(f^0, f)(\overline{u}), \\ z \in N_E(\overline{e}). \end{cases} \tag{6}$$

• Similar results can be found in the books:

H. O. Fattorini (1999),

J. F. Bonnans & A. Shapiro (2000).

• The point is how to verify the assumption **(H)** for concrete problems.

3. Characterization of finite codimensionality

$\diamondsuit$ Some reductions

In order to guarantee the assumption $(\mathbf{H})$ in Theorem 1 to hold, it suffices to prove the finite codimensionality of the following set M in X:

$$M = \{\, f'(\overline{u})v \;\big|\; |v|_V \leq 1 \,\}. \tag{7}$$

Indeed, $M \subseteq$ Var $f(\overline{u})$ and the finite codimensionality of Var $f(\overline{u})$ in X implies that of Var $f(\overline{u}) - I_E(\overline{e}) \cap B_\rho(0)$.

Put

$$N = \{\, f'(\overline{u})v \mid v \in V \,\}.$$

Then

$$\text{span} M = \overline{N} \quad \text{and} \quad \overline{\text{co}} M = M. \tag{8}$$

Clearly, the set M is finite codimensional iff so is the set N.

$\diamondsuit$ Duality argument:

• The finite codimensionality of N is equivalent to a suitable estimate.

Theorem 2: *The following two assertions are equivalent:*

(1) *The set* N *is a finite codimensional subspace in* X.

(2) $\exists$ *a finite codimensional subspace* $\widetilde{X}$ *of* X', *such that*

$$|\phi|_{X'} \leq C|f'(\overline{u})^*\phi|_{V'}, \qquad \forall\, \phi \in \widetilde{X}. \tag{9}$$

• In general, it is hard to find the finite codimensional subspace $\widetilde{X}$ of X' in the assertion **(2)** of Theorem 2.

• We give below another equivalent criterion, where the estimate holds on the whole space X':

Theorem 3: *The following two assertions are equivalent:*
(1) $\exists$ *a compact operator* G *from* X' *to a Banach space* W*, such that*

$$|\phi|_{X'} \leq C\big(|f'(\overline{u})^*\phi|_{V'} + |G\phi|_W\big), \qquad \forall\, \phi \in X'. \tag{10}$$

(2) $\exists$ *a finite codimensional subspace* $\widetilde{X}$ *of* X'*, such that*

$$|\phi|_{X'} \leq C|f'(\overline{u})^*\phi|_{V'}, \qquad \forall\, \phi \in \widetilde{X}. \tag{11}$$

• A key observation: The estimate (17) is ROBUST with respect to compact perturbations!

4. Application to deterministic optimal control problems

Let Y and U be two Hilbert spaces. Consider the following evolution equation:

$$y_t(t) = Ay(t) + f(t, y(t), u(t)) \quad t \in [0, T], \tag{12}$$

where $T > 0$, u is the control variable valued in U, y is the state variable valued in Y, A generates a C_0-semigroup in Y, and f satisfies suitable conditions.

Let $\widetilde{U}$ be a nonempty subset of U,

$$\mathcal{U} = \left\{ u \in L^2(0, T; U) \;\middle|\; u : (0, T) \to \widetilde{U} \text{ is measurable} \right\},$$

and S be a subset of $Y \times Y$.

For a suitable real function f^0, define the set of admissible pairs:

$$\mathcal{A}_{ad} = \Big\{ (u, y) \in \mathcal{U} \times C([0,T];Y) \;\Big|\; (y(0), y(T)) \in S, \\ y \text{ is a mild solution of (19) associated to } u \\ \text{and } f^0(\cdot, y(\cdot), u(\cdot)) \in L^1(0,T) \Big\},$$

and the cost functional:

$$J(u(\cdot), y(\cdot)) = \int_0^T f^0(t, y(t), u(t))dt, \quad \forall\, (u, y) \in \mathcal{A}_{ad}.$$

Usually, one assumes that $\mathcal{A}_{ad} \neq \emptyset$. The optimal control problem considered here is stated as follows:

(P) Find $(\overline{u}, \overline{y}) \in \mathcal{A}_{ad}$ so that $J(\overline{u}, \overline{y}) = \inf_{(u,y) \in \mathcal{A}_{ad}} J(u, y)$.

Pontryagin type maximum principle: Let $(\overline{u}, \overline{y})$ be an optimal pair. Then $\exists$ a pair of $(\psi^0, \psi(\cdot)) \in \mathbb{R} \times C([0,T]; Y')$, so that

$$|\psi^0|^2 + |\psi(t)|^2_{Y'} > 0, \ \forall\, t \in [0,T];$$

$$\psi_t = -A^*\psi - f_y(t, \overline{y}(t), \overline{u}(t))^*\psi - \psi^0 f^0_y(t, \overline{y}(t), \overline{u}(t));$$

$$\langle \psi(0), y^0 - \overline{y}(0) \rangle_{Y',Y} - \langle \psi(T), y^1 - \overline{y}(T) \rangle_{Y',Y} \le 0, \forall (y^0, y^1) \in S;$$

$$H(t, \overline{y}(t), \overline{u}(t), \psi^0, \psi(t)) = \max_{u \in \widetilde{U}} H(t, \overline{y}(t), u, \psi^0, \psi(t)),$$

$$\text{a.e. } t \in (0,T),$$

where $H(t, y, u, \psi^0, \psi) = \psi^0 f^0(t, y, u) + \langle \psi, f(t, y, u) \rangle_{Y',Y}$.

In order to guarantee the nontriviality of $(\psi^0, \psi(\cdot))$, one needs some sort of finite codimensionality condition. For this, consider the controlled system:

$$\begin{cases} \xi_t = A\xi + f_y(t, \overline{y}(t), \overline{u}(t))\xi \\ \qquad\qquad + \Big[f(t, \overline{y}(t), u(t)) - f(t, \overline{y}(t), \overline{u}(t)) \Big], \ t \in (0, T] \\ \xi(0) = 0, \end{cases} \quad (13)$$

and the set (in Y)

$$\mathcal{R} = \Big\{ \xi(T) \ \Big| \ \xi \text{ is the solution of (13) for some } u(\cdot) \in \mathcal{U} \Big\}.$$

Also, consider the homogenous equation:

$$\begin{cases} \eta_t = A\eta + f_y(t, \overline{y}(t), \overline{u}(t))\eta \quad t \in [0, T], \\ \eta(0) = y^0, \end{cases} \tag{14}$$

and another set (in Y):

$$\mathcal{Q} = \left\{ y^1 - \eta(T) \,\middle|\, \eta \text{ is the solution of } (14) \text{ and } (y^0, y^1) \in S \right\}.$$

Theorem (X. Li-J. Yong, 1991): If $\mathcal{R} - \mathcal{Q}$ is finite codimensional in Y, then the optimal pair $(\overline{u}, \overline{y})$ satisfies Pontryagin type maximum principle.

$\diamondsuit$ Reduction of the finite codimensionality condition to suitable *a priori* estimates.

Consider the linear controlled system:

$$\begin{cases} y_t = Ay(t) + F(t)y(t) + B(t)u(t) \qquad t \in [0, T], \\ y(0) = y_0, \end{cases} \tag{15}$$

where $F(\cdot) \in L^\infty(0, T; \mathcal{L}(Y))$, $B(\cdot) \in L^\infty(0, T; \mathcal{L}(U; Y))$, $y_0 \in Y$ and $u \in L^2(0, T; U)$.

Let $\widetilde{\mathcal{U}}$ be a nonempty bounded set of $L^2(0, T; U)$ and $\overline{\text{co}}\,\widetilde{\mathcal{U}}$ have at least an interior point. Also, set

$$M = \Big\{ y(T; 0, u) \in Y \;\Big|\; y \text{ is the solution of } (15) \text{ with } y_0 = 0 \text{ and some } u(\cdot) \in \widetilde{\mathcal{U}} \Big\}.$$

Theorem 4: *The set M is finite codimensional in Y iff (15) is finite codimensional exactly controllable.*

Consider the dual system of (15):

$$\begin{cases} \phi_t(t) = -A^*\phi(t) - (F(t))^*\phi(t) \qquad t \in [0,T], \\ \phi(T) = \phi_T, \end{cases} \tag{16}$$

where $\phi_T \in Y'$.

Theorem 5: *The following assertions are equivalent:*

(1) *The system* (15) *is finite codimensional exactly controllable.*

(2) *There is a compact operator* $\mathcal{G}$ *from* Y' *to a Banach space* Z*, so that any solution* ϕ *of* (16) *satisfies that*

$$|\phi_T|_{Y'} \le C\Big[|(B(\cdot))^*\phi|_{L^2(0,T;U)} + |\mathcal{G}\phi_T|_Z\Big], \quad \forall\, \phi_T \in Y'. \tag{17}$$

Remark. The inequality (17) can be regarded as a Gårding type inequality, which concerns the lower bound of a bilinear form induced by a linear elliptic (pseudo-)differential operator.

To see this, let us recall the classical Gårding inequality: Let Ω be a bounded domain in $\mathbb{R}^n$ with a smooth boundary Γ, and L a uniformly linear elliptic differential operator of order $2k$ (for positive integers n and k). Then, for any $v \in H_0^k(\Omega)$ and some constants $C_0 > 0$ and $C_1 \geq 0$,

$$C_0 |v|^2_{H^k(\Omega)} \leq \langle Lv, v \rangle_{H^{-k}(\Omega), H_0^k(\Omega)} + C_1 |v|^2_{L^2(\Omega)}. \tag{18}$$

Clearly, both (17) and (18) have an extra term, that is, $|G\phi_T|_X$ and $|v|^2_{L^2(\Omega)}$, respectively. Also, these two terms are accordingly compact with respect to the ones in the left hand sides of the corresponding estimates.

In summary:

The finite codimensionality condition

in our optimal control problem with endpoint constraints

$\longrightarrow$ Finite codimensional exact controllability

$\longrightarrow$ *A priori* estimate for the dual equation.

How to derive the above *a priori* estimate is purely a PDE problem, which can be solved for some concrete situations.

• Again: The above *a priori* estimate is ROBUST with respect to compact perturbations! Note however that exact controllability does NOT enjoy the same property.

5. Application to stochastic optimal control problems

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete filtered probability space, on which a one-dimensional standard Brownian motion $\{B(t)\}_{t\geq 0}$ is defined, such that $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration generated by $B(\cdot)$, augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$.

Consider the following stochastic differential equation:

$$\begin{cases} dx(t) = a(t, x(t), u(t))dt + b(t, x(t), u(t))dB(t), \\ x(0) = x_0, \end{cases} \tag{19}$$

where $a, b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \mathbb{R}^n$ satisfy suitable conditions, so that for any $x_0 \in \mathbb{R}^n$ and $u \in \mathcal{U} = L^2_{\mathbb{F}}(0, T; \mathbb{R}^m)$, (19) is well-posed.

Write

$$J(u(\cdot)) = \mathbb{E}\int_0^T f^0(t, x(t), u(t))dt, \quad \forall\, u(\cdot) \in \mathcal{U},$$

where $x(\cdot)$ is the solution of (19) associated to $u(\cdot)$. Assume that $f^0 : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ satisfies suitable conditions, so that for any $u(\cdot) \in \mathcal{U}$, $f^0(\cdot, x(\cdot), u(\cdot)) \in L^1_{\mathbb{F}}(0, T)$. For any given $x_1 \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n)$, set

$$\mathcal{U}_{ad} = \left\{ u(\cdot) \in \mathcal{U} \;\middle|\; x(T) = x_1 \right\}.$$

Suppose that $\overline{u}(\cdot) \in \mathcal{U}$ is an optimal control of the following optimal control problem:

$$J(\overline{u}(\cdot)) = \inf_{u(\cdot)\in\mathcal{U}_{ad}} J(u(\cdot)).$$

Denote by $\overline{x}$ the solution of (19) associated to $\overline{u}$.

In order to derive the first order necessary condition for the above $\overline{u}$, put

$$A_1(t) = a_x(t, \overline{x}(t), \overline{u}(t)), \quad C_1(t) = a_u(t, \overline{x}(t), \overline{u}(t)),$$
$$A_2(t) = b_x(t, \overline{x}(t), \overline{u}(t)), \quad C_2(t) = b_u(t, \overline{x}(t), \overline{u}(t)),$$

and consider the following stochastic differential equation:

$$\begin{cases} dz = (A_1(t)z + C_1(t)v)\, dt + (A_2(t)z + C_2(t)v)\, dB(t), \\ z(0) = 0, \end{cases} \tag{20}$$

where $v \in \mathcal{U}$. In order to guarantee the necessary condition to be nontrivial, it suffices that the following set is finite codi-mensional:

$$M = \left\{ z(T) \in L^2_{\mathcal{F}_T}(\Omega; \mathbb{R}^n) \;\middle|\; |v|_{\mathcal{U}} \le 1 \right\}.$$

By duality, we consider the following backward stochastic differential equation:

$$\begin{cases} d\phi = -\left(A_1^\top(t)\phi + A_2^\top(t)\Phi\right)dt + \Phi dB(t), \\ \phi(T) = \phi_T. \end{cases} \tag{21}$$

Then M is finite co-dimensional is equivalent to the following inequality holds for solutions of (21):

$$\mathbb{E}|\phi_T|^2_{\mathbb{R}^n} \le C\Big[\mathbb{E}\int_0^T |C_1^\top(t)\phi(t) + C_2^\top(t)\Phi(t)|^2_{\mathbb{R}^m}dt + |G\phi_T|^2_Y\Big], \\ \forall\ \phi_T \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n),$$

for some compact operator G from $L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$ to a Banach space Y.

• Once more: The above estimate is ROBUST with respect to compact perturbations!

Thank you！