数理解析研究所講究録
1187 巻 2001 年 58-65

Conjugate points for a nonlinear programming problem [1]

Hidefumi KAWASAKI
Graduate School of Math., Kyushu Univ.

Nonlinear Analysis and Convex Analysis,
Organizer: Prof. W. Takahashi
Aug. 29, 2000,

"**Conjugate points**" is a global concept in calculus of variation, and plays an important role in discussing optimality. Though it has been a theme of differential geometry, mathematical programming approach has been recently developed with several extensions of the conjugate points theory to optimal control problems and variational problems with state constraints. In these extremal problems, the variable is not a vector x in R^n but a function $x(t)$. So a simple and natural question arises. Is it possible to establish a conjugate points theory for a minimization problem: minimize $f(x)$ on $x \in R^n$? In [3], the author positively answered this question. He introduced "the Jacobi equation" and "conjugate points" for it, and describe optimality conditions in terms of "conjugate points". In this paper, we extend it to a constrained nonlinear programming.

1. Jacobi's conjugate points theory

In this section, we review the classical conjugate points theory for the simplest problem in calculus of variations :

$$(SP) \quad \text{Minimize} \quad \int_0^T f(t, x(t), \dot{x}(t))dt$$
$$\text{subject to} \quad x(0) = A, \; x(T) = B.$$

If $\bar{x}$ is a weak minimum for (SP), then it satisfies the Euler equation $df_{\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t))/dt = f_x(t, \bar{x}(t), \dot{\bar{x}}(t))$ and the Legendre condition $f_{\dot{x}\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t)) \geq 0$. Legendre attempted to prove its inverse, that is, he expected that if a feasible solution $\bar{x}(t)$ satisfies the Euler equation and the strengthened Legendre condition: $f_{\dot{x}\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t)) > 0$, then $\bar{x}(t)$ would be a weak minimum. However, his conjecture was false. Jacobi solved this problem by introducing "conjugate points" in 1837. Let $\bar{x}(t)$ be a feasible solution for (SP), then conjugate points are defined via the Jacobi equation:

$$\frac{d}{dt}\{\bar{f}_{\dot{x}x}(t)y(t) + \bar{f}_{\dot{x}\dot{x}}(t)\dot{y}(t)\} = \bar{f}_{xx}(t)y(t) + \bar{f}_{x\dot{x}}(t)\dot{y}(t). \tag{1}$$

where $\bar{f}_{\dot{x}\dot{x}}(t) := f_{\dot{x}\dot{x}}(t, \bar{x}(t), \dot{\bar{x}}(t))$, etc. A point $c \in (0, T]$ is said to be *conjugate to* $t = 0$ if there exists a non-trivial solution $y(t)$ of the Jacobi equation (1) on $[0, c]$ and $y(0) = y(c) = 0$. Then Jacobi proved the following result:

THEOREM 1 *(Jacobi) If $\bar{x}(t)$ is a weak minimum for the simplest problem (SP) and it satisfies the strengthened Legendre condition, then there is no point conjugate to $t = 0$ on $[0, T)$. Conversely, if $\bar{x}(t)$ satisfies the Euler equation and the strengthened Legendre condition, and if there is no point conjugate to $t = 0$ on $[0, T]$, then $\bar{x}(t)$ is a weak minimum.*

[1] Supported in part by Grant-in-Aid for Scientific Research No.11440033 from the Ministry of Education, Science, Sports, and Culture of Japan.

Recently, "conjugate points" has been extended to complex extremal problems such as optimal control problems and variational problems with state constraints, see e.g. Kawasaki-Zeidan [5], Loewen-Zheng [7], Warga [9], Zeidan [10], and Zeidan-Zezza [11][12][13]. The present paper is outside of this trend. In this paper, we first quote the conjugate points theory [3] for an unconstrained nonlinear programming problem (P):

$$(P) \qquad \text{Minimize} \quad f(x) \quad \text{on} \quad x \in R^n,$$

where $f : R^n \to R$ is assumed to be twice continuously differentiable. It seems to the author that one reason why researchers have not paid much attention to conjugate points for (P) lies in the following elementary results on it:

THEOREM 2 *If $\bar{x}$ is a minimum for (P), then it satisfies $f'(\bar{x}) = 0$ and $f''(\bar{x}) \geq 0$, Conversely, if $\bar{x}$ satisfies $f'(\bar{x}) = 0$ and $f''(\bar{x}) > 0$, then it is a minimum for (P), where "$\geq$" and "$>$" stand for non-negative definite and positive definite, respectively.*

Theorem 2 seems to assert that there is no room for "conjugate points" to play a role in (P). However, an interesting connection between Jacobi's condition and the theory of quadratic forms was discussed in Gelfand and Fomin [1, p.125]. We quote it in the next section. Furthermore, the author [2] recently found a stimulating example that strongly indicates possibility to establish a conjugate points theory for (P), see Example 1 below.
Example 1 ([2]) Let us first consider the shortest path problem on the unit sphere S in R^3. The problem is finding a shortest path on S joining two points $A = (1, 0, 0)$ and $B = (\cos T, \sin T, 0)$, where $0 < T < 2\pi$ is given. When $T > \pi$, the equatorial arc, say AB, in Fig. 1 is not a weak minimum. Indeed, take another great circle arc (the broken curve in Fig. 1) joining A and $C = (-1, 0, 0)$, say AC, and join AC and the equatorial arc CB. Then it has the same length with the equatorial arc AB. However, we get a shorter curve by taking a short cut around C.

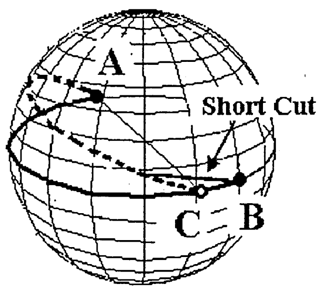

Figure 1

According to the classical conjugate point theory, C is conjugate to A. Next, let us approximate the shortest path problem by an extremal problem in a finite dimensional space as follows:

1: Take a finite number of equally located longitudes $\ell_0, \ \ell_1, \ldots, \ell_{n+1}$, where we assume that $A \in \ell_0$ and $B \in \ell_{n+1}$, see Fig. 2.

2: Choose one point, say X_k, on each ℓ_k for $k = 1, 2, \ldots n$.

3: Minimize the length of the polygonal curve joining $A, X_1, \ldots, X_n, B$.

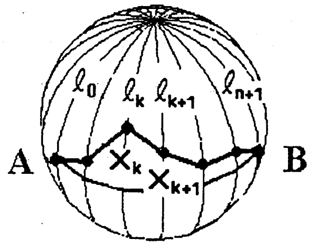

Figure 2

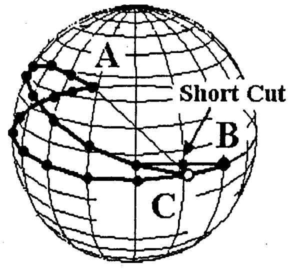

Figure 3

Then the above observation on the classical shortest path problem can be also applied to this extremal problem in R^n. In fact, by taking a short cut around C in Fig. 3, we obtain a shorter polygonal curve. Hence C can be regarded as a conjugate point in the finite-dimensional analogue.

2. Connection between Jacobi's condition and Quadratic Form

In this section, we quote from [1, p.127] an interesting observation on a connection between Jacobi's condition and the theory of quadratic forms in R^n. According to the classical conjugate point theory, the quadratic functional

$$\int_0^T \{P\dot{y}^2 + Ry^2\}dt, \tag{2}$$

where $P(t) > 0$ and $R(t) \in R$, is positive for all non-trivial $y(t)$ such that $y(0) = y(T) = 0$ if and only if $[0, T]$ contains no point conjugate to 0, where the corresponding Jacobi equation is

$$\frac{d}{dt}\{P\dot{y}\} = Ry,$$

see [1, p.111]. By introducing the points $0 = t_0 < t_1 < \cdots < t_n < t_{n+1} = T$, we get $n+1$-equal parts of length $\Delta t := T/(n+1)$. Then the quadratic functional (2) is approximated by the quadratic form:

$$\sum_{k=0}^{n} \left\{ P_k \left(\frac{y_{k+1} - y_k}{\Delta t} \right)^2 + R_k y_k^2 \right\} \Delta t, \tag{3}$$

where $P_k := P(t_k)$, $R_k := R(t_k)$, $y_k := y(t_k)$, and $y_0 = y_{n+1} = 0$. By putting $a_k := R_k\Delta t + (P_{k-1} + P_k)/\Delta t$, $b_k := -P_{k-1}/\Delta t$, $y = (y_1, \ldots, y_n)$, and

$$A_n := \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & a_n \end{pmatrix}, \tag{4}$$

quadratic form (3) is expressed as $y^T A_n y$. So, let us now discuss positive-definiteness of A_n. According to Sylvester's criterion, A_n is positive definite if and only if its descending principal minors $|A_k|$, $(k = 1, \dots, n)$ are positive. On the other hand, there is a recurrence formula for tribunal matrix (4)

$$|A_k| = a_k |A_{k-1}| - b_{k-1}^2 |A_{k-2}|. \tag{5}$$

By making the change of variables $Y_0 := 0$, $Y_1 := \Delta t$, and

$$Y_{k+1} := \frac{(\Delta t)^{k+1} |A_k|}{P_1 \cdots P_k}, \quad k = 1, \dots, n,$$

recursion relation (5) reduces to

$$\frac{P_k \frac{Y_{i+1} - Y_k}{\Delta t} - P_{k-1} \frac{Y_k - Y_{k-1}}{\Delta t}}{\Delta t} = R_k Y_k. \tag{6}$$

Tending $\Delta t \to 0$ in (6), we get the Jacobi equation $d(P\dot{Y})/dt = RY$. Therefore expansion (5) can be regarded as the Jacobi equation.

3. Conjugate Points for (P)

In this section, we define the Jacobi equation and conjugate points for (P), and describe sufficient optimality conditions in terms of conjugate points.

First, let us discuss positive-definiteness of symmetric matrices. According to Sylvester's criterion, an $n \times n$-symmetric matrix $A = (a_{ij})$ is positive-definite if and only if its descending principal minors $|A_k|$ $(k = 1, \dots, n)$ are positive, where

$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}.$$

The following lemma shows that the determinant of any square matrix is expanded w.r.t. the descending principal minors:

LEMMA 1 *([3]) For any $n \times n$-matrix $A = (a_{ij})$, its determinant is expanded as follows:*

$$|A| = \sum_{k=0}^{n-1} \sum_{\rho} \varepsilon(\rho) a_{k+1\rho(k+1)} a_{k+2\rho(k+2)} \cdots a_{n\rho(n)} |A_k| \tag{7}$$

where $|A_0| := 1$, $\varepsilon(\rho)$ denotes the sign of ρ, and the summation is taken over all permutations ρ on $\{k+1, \dots, n\}$ satisfying that there is no $\ell > k$ such that ρ is closed on $\{\ell + 1, \dots, n\}$.

When A is tridiagonal, expansion (7) reduces to (5). So we obtain our definition of the Jacobi equation and conjugate points for (P):

DEFINITION 1 *([3]) For any $n \times n$-matrix $A = (a_{ij})$, we call the recursion relation on $y_0, \dots, y_n$*

$$y_k = \sum_{i=0}^{k-1} \sum_{\rho} \varepsilon(\rho) a_{i+1\rho(i+1)} a_{i+2\rho(i+2)} \cdots a_{k\rho(k)} y_i, \qquad k = 1, \dots, n \tag{8}$$

the Jacobi equation for A. We say that k is conjugate to 1 if the solution $\{y_i\}$ of the Jacobi equation with $y_0 > 0$ changes the sign from positive to non-positive at k. Namely,

$$y_0 > 0,\ y_1 > 0,\ \dots, y_{k-1} > 0, \ \text{and}\ y_k \le 0. \tag{9}$$

THEOREM 3 *([3]) For any (n, n)-symmetric matrix A, A is positive-definite if and only is there is no point conjugate to 1.*

THEOREM 4 *([3]) A sufficient condition for an extremal $\bar{x}$ to be a minimum for (P) is that, concerning the Hesse matrix $f''(\bar{x})$, there is no point conjugate to 1.*

REMARK 1 *It is easy to give a necessary optimality condition for (P) in terms of conjugate points, see [3]. However, since the descending principal minors $|A_1|, \dots, |A_n|$ are not enough to characterize $A \ge 0$, the situation is slightly different from the sufficiency case. Namely, though the implication $A \ge 0 \Rightarrow |A_k| \ge 0$ $(1 \le k \le n)$ is true, its inverse is not true in general. In order to check $A \ge 0$, we have to test the sign of all principal minors.*

4. Example

Example 2 ([3]) This example is same with Example 1. Here we compute conjugate points for it. By means of the spherical coordinates, any point on the k-th longitude ℓ_k is expressed as $X_k = (\sin\theta(t_k)\cos t_k, \sin\theta(t_k)\sin t_k, \cos\theta(t_k))$. Hence the minimization problem of the length of the polygonal arc joining $n+2$ points $A = X_0,\ X_1,\ \dots,\ X_{n+1} = B$ is formulated as follows:

$$(P_1) \quad \text{Minimize} \quad f(\theta_1, \dots, \theta_n) := \sum_{k=0}^{n} \sqrt{2(1 - \cos\Delta t \sin\theta_{k+1}\sin\theta_k - \cos\theta_{k+1}\cos\theta_k)},$$

where $\theta_0 = \theta_{n+1} = \pi/2$. The variable $\theta \in R^n$ that corresponds to the equatorial arc $\bar{\theta}(t) \equiv \pi/2$ is $\bar{\theta} := (\pi/2, \dots, \pi/2)$. Then the Hesse matrix of f at $\bar{\theta}$ is

$$f''(\bar{\theta}) = \frac{1}{\sqrt{2(1-c)}} \begin{pmatrix} 2c & -1 & & \\ -1 & 2c & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2c \end{pmatrix} \tag{10}$$

where $c := \cos\Delta t$. Since

$$k \left\{ \begin{vmatrix} 2c & -1 & & \\ -1 & 2c & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2c \end{vmatrix} \right. = \frac{\sin(k+1)\Delta t}{\sin\Delta t},$$

the principle minor of size k of (10) is positive if $(k+1)\Delta t < \pi$. Since $\Delta t = T/(n+1)$, we conclude that

(a) when $T < \pi$, there is no point conjugate to 1,
(b) when $T \geq \pi$, the first number k satisfying $\frac{k+1}{n+1} \geq \pi/T$ is conjugate to 1,

which matches the classical result.

5. Constrained case

In this section, we deal with a nonlinear programming problem with equality constraints:

$$(P_2) \qquad \begin{array}{ll} \text{Minimize} & f(x) \\ \text{subject to} & h(x) := (h_1(t), \ldots, h_m(x))^T = (0, \ldots, 0)^T, \\ & x \in R^n \end{array}$$

Before going further, we remark that all the results in this section can be easily extended to the case where there are equality and inequality constraints under the strict complimentarity assumption.

Now, let $\bar{x}$ be a feasible solution for (P_2). In the following, we assume that $h_j'(\bar{x})$, $(j \in J)$ are linearly independent. Then there exists $\lambda \in R^m$ such that $L'(\bar{x}) = 0$ and

$$\xi^T L''(\bar{x})\xi \geq 0 \quad \text{if} \quad h'(\bar{x})\xi = 0, \tag{11}$$

where $L(x) := f(x) + \lambda^T h(x)$. Conversely, if there exists $\lambda \in R^m$ such that $L'(\bar{x}) = 0$ and

$$\xi^T L''(\bar{x})\xi > 0 \quad \text{if} \quad h'(\bar{x})\xi = 0 \ (\xi \neq 0), \tag{12}$$

then $\bar{x}$ is a minimum, see Fiacco and McCormick.

Since $h'(\bar{x})$ has full rank, $m \times n$-matrix $h'(\bar{x})$ can be divided as $h'(\bar{x}) = (B, N)$, where B is an $m \times m$-nonsingular matrix and N is an $m \times (n-m)$-matrix. Similarly, we divide $\xi \in R^n$ as $\xi = (y, z) \in R^m \times R^{n-m}$. Then $h'(\bar{x})\xi = 0$ is equivalent to $y = -B^{-1}Nz$, so that

$$\xi^T L''(\bar{x})\xi = z^T(-N^T B^{-T}, I)L''(\bar{x}) \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} z \tag{13}$$

Hence the following matrix is the key to describe optimality conditions.

$$K := (-N^T B^{-T}, I)L''(\bar{x}) \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} \tag{14}$$

THEOREM 5 *A sufficient condition for a feasible solution $\bar{x}$ be a minimum for (P_2) is that there exists $\lambda \in R^m$ such that $L'(\bar{x}) = 0$ and, concerning K defined by (14), there is no point conjugate 1.*

Example 3 In Example 2, we computed conjugate points for a finite-dimensional analogue to the classical shortest path problem on the unit sphere S. Here we present another approach. It is treating the problem as a constrained problem:

1. For $k = 1, \ldots, n+1$, let R_k be the ring defined by $\{(\cos \frac{kT}{n+1}, y, z) \in S\}$.
2: Choose one point, say X_k, on each R_k for $k = 1, 2, \ldots n$.
3: Minimize the length of the polygonal curve joining $A = (1,0,0), X_1, \ldots, X_n, B = (\cos T, \sin T, 0)$.

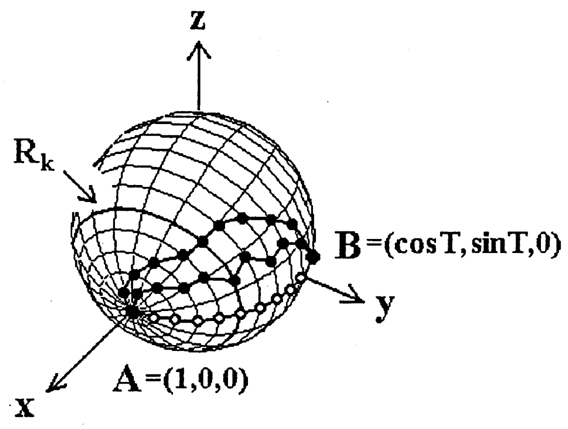

Figure 4

Then this finite-dimensional analogue is formulated as follows:

$$\begin{aligned}
\min \quad & f(y_1,\ldots,y_n,z_1,\ldots,z_n) \\
& := \sum_{k=0}^{n} \sqrt{(y_{k+1}-y_k)^2+(z_{k+1}-z_k)^2+(\cos(k+1)\Delta t-\cos k\Delta t)^2} \\
\text{s.t.} \quad & (y_0,z_0)=(0,0), \quad (y_{n+1},z_{n+1})=(\sin T,0) \\
& h_k(y_1,\ldots,y_n,z_1,\ldots,z_n) := y_k^2+z_k^2-\sin^2 k\Delta t=0, \quad k=1,\ldots,n,
\end{aligned}$$

where $\Delta t := T/(n+1)$. Since the equatorial arc corresponds to

$$(\bar{y},\bar{z}) := (\bar{y}_1,\ldots,\bar{y}_n,\bar{z}_1,\ldots,\bar{z}_n) = (\sin\Delta t,\ldots,\sin n\Delta t,0,\ldots,0)$$

it can be easily seen that

$$h'(\bar{y},\bar{z}) = 2\begin{pmatrix} \sin\Delta t & & & & 0 & \cdots & 0 \\ & \sin 2\Delta t & & & 0 & \cdots & 0 \\ & & \ddots & & \vdots & & \vdots \\ & & & \sin n\Delta t & 0 & \cdots & 0 \end{pmatrix} \tag{15}$$

Hence we may choose as B and N in (13)

$$B = 2\begin{pmatrix} \sin\Delta t & & & \\ & \sin 2\Delta t & & \\ & & \ddots & \\ & & & \sin n\Delta t \end{pmatrix}, \qquad N = \begin{pmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

respectively. Furthermore, we get $L(y,z) = f(y,z) - \sin\Delta t\sum_{k=1}^{n}(y_k^2+z_k^2-\sin^2 k\Delta t)$. Hence

$$(-N^TB^{-T},I)\begin{pmatrix} L_{yy} & L_{yz} \\ L_{xx} & L_{zz} \end{pmatrix}\begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} = L_{zz} = \frac{1}{\sin\frac{\Delta t}{2}}\begin{pmatrix} 2c & -1 & & \\ -1 & 2c & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2c \end{pmatrix}$$

where $c := \cos \Delta t$, which is same with (10) up to constant. Therefore we get the same conclusion with Example 2.

References

[1] I. M. Gelfand and S. V. Fomin, *Calculus of variations,* Prentice Hall, 1963.

[2] H. Kawasaki, Conjugate points for a minimization problem in a finite dimensional space, Kyushu University Preprint Series in Mathematics, No. 23 (1999), pp. 1-12.

[3] H. Kawasaki, A conjugate points for a minimization problem in a finite dimensional space, revised.

[4] 川崎英文,「極値問題の大域的性質：共役点」, 第11回RAMPシンポジウム論文集, 85-96 (1999, Oct.19).

[5] H. Kawasaki and V. Zeidan, Conjugate points for variational problems with equality and inequality state constraints, to appear in SIAM J. Control Optim.

[6] S. Koga and H. Kawasaki, Legendre-type optimality conditions for a variational problem with inequality state constraints, Math. Programming, 84 (1999), pp. 421–434.

[7] P. D. Loewen and H. Zheng, Generalized conjugate points for optimal control problems, Nonlinear Anal., 22 (1994), pp. 771–791.

[8] G. Strang, *Linear Algebra and its Applications,* Academic Press, New York, 1976.

[9] J. Warga, A second-order Lagrangian condition for restricted control problems, J. Optim. Theory Appl., 24 (1978), pp. 475-483.

[10] V. Zeidan, The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency, SIAM J. Control and Optim., 32 (1994), pp. 1297-1321.

[11] V. Zeidan and P. Zezza, The conjugate point condition for smooth control sets, J. Math. Anal. Appl., 132 (1988), pp. 572-589.

[12] V. Zeidan and P. Zezza, Coupled points in the calculus of variations and applications to periodic problems, Trans. Amer. Math. Soc., 315 (1989), pp. 323-335.

[13] V. Zeidan and P. Zezza, Coupled points in optimal control theory, IEEE Trans. Automat. Control, 36 (1991), pp. 1276-1281.

kawasaki@math.kyushu-u.ac.jp