Transversality Theorem : A Useful Tool for Establishing Genericity

S. S. Keerthi, N. K. Sancheti and A. Dattasharma

Department of Computer Science & Automation
Indian Institute of Science, Bangalore-560 012, India

Abstract

It is shown, via a number of examples in linear algebra and control, optimization, and geometry, that the transversality theorem of differential topology is a useful tool for establishing genericity of a property which is a function of a finite number of real parameters.

1 Introduction

Consider $\mathcal{P}(y)$, a property $\mathcal{P}$ which is a function of y , a real finite-dimensional vector of parameters. Let y take values in Y , an open set. We will say, "property $\mathcal{P}(y)$ holds for a.a. $y \in Y$ " ('a.a' denoting 'almost all') to mean that "the set, $\{y \in Y : \mathcal{P}(y) \text{ does not hold}\}$ has zero (Lebesgue) measure". If $\mathcal{P}(y)$ holds for a.a. $y \in Y$ then we say that $\mathcal{P}$ is *generic* in Y .

Hirsch and Smale [6] give a definition of genericity which is different from ours. They call a property as generic if $\tilde{Y} = \{y \in Y : \mathcal{P}(y) \text{ holds}\}$ contain an open dense set. It is easy to give examples to show that the two definitions are different, i.e., one does not imply the other. If a property is generic by our definition then $\tilde{Y}$ is dense because $Y \setminus \tilde{Y}$ is of measure zero; however $\tilde{Y}$ may not contain an open dense set. On the other hand, if $\tilde{Y}$ contains an open dense set, it does not imply that $Y \setminus \tilde{Y}$ has zero measure. It appears that our definition is more commonly used in the literature. Both definitions are reasonable. All references to genericity in this paper correspond to our definition.

Suppose X , Y and Z are open sets in some finite-dimensional real spaces, Z contains the origin, and there exists a differentiable function $F : X \times Y \rightarrow Z$ such that, for any given $y \in Y$, $\mathcal{P}(y)$ holds iff $F(x, y) = 0$ has (or, does not have) a solution, $x \in X$. In such a situation, transversality theorem [5], or, parametrized Sard's theorem [10] as it is otherwise called, is a very useful tool for establishing that $\mathcal{P}$ is generic in Y . The aim of this paper is to illustrate this point via a number of examples in linear algebra and control, optimization, and geometry. In many of these examples the genericity of the property involved is intuitively obvious and yet, establishing genericity by direct arguments is hard. In such situations the transversality theorem is a very useful tool.

The organization of the paper is as follows. In section 2 we state a simple version of the transversality theorem which suffices for all our examples. Examples from linear algebra and control, optimization, and geometry are given in sections 3, 4 and 5, respectively. Section 6 contains some concluding remarks.

We end this section with a list of some notations used in the paper and an useful lemma on Lebesgue measure. E^n will denote the n -dimensional Euclidean real space. If $x \in E^n$ then x_i = the i -th component of x , x^T = transpose of x , and $\text{diag}(x) = n \times n$ diagonal matrix whose i -th diagonal element is x_i , $i = 1, \dots, n$. If $\{x^1, x^2, \dots, x^k\}$ is a sequence of vectors in E^n then $(x^1, x^2, \dots, x^k)$ will denote the vector, $((x^1)^T | (x^2)^T | \dots | (x^k)^T)^T \in E^{kn}$. I_n is the $n \times n$ identity matrix and $\det(A)$ = determinant of the square matrix, A . C^k will denote the class of k -times continuously differentiable functions. If $F : A \rightarrow B$ where A and B are open in E^k and E^l respectively, and $F \in C^1$, then, for $\alpha \in A$, $JF(\alpha)$ will denote the Jacobian of F at α ; $JF(\alpha)$ is an $l \times k$ matrix. Also, if $\alpha_1 \in E^{k_1}$ is a subvector of α formed by choosing some specified components of α , then $F_{\alpha_1}(\alpha)$ will denote the $l \times k_1$ matrix of partial derivatives of F with respect to the variables in α_1 evaluated at α . In particular, $JF(\alpha) = F_{\alpha}(\alpha)$. If $f : E^n \rightarrow E$, $\nabla f(x)$ and f_{xx} will respectively denote the $n \times 1$ gradient and the $n \times n$ Hessian of f evaluated at x . ϕ will denote the null set.

The following well-known result [11] is useful.

Lemma 1.1. (a) A finite union of measure zero sets has zero measure. (b) If X and Y are open sets in some finite dimensional real spaces, $\tilde{X} \subset X$ and $\tilde{X}$ has zero measure, then $\tilde{X} \times Y$ has zero measure.

2 Transversality Theorem

Consider a function $F : X \times Y \rightarrow E^N$ where X and Y are open sets in E^L and E^M respectively. Let $F \in C^1$. We say that $(x, y) \in X \times Y$ is a *regular point* of F if $JF(x, y)$ has full row rank. Let 0 denote the origin in E^N . We say that 0 is a *regular value* of F if (x, y) is a regular point $\forall (x, y) \in F^{-1}(0) = \{(x, y) : F(x, y) = 0\}$. A weak version of the transversality theorem of differential topology [5], one which is sufficient for all our needs in this paper, is the following.

Transversality Theorem. Let $F : X \times Y \rightarrow E^N$ be as given above and 0 be a regular value of F . Then 0 is a regular value of $F(\cdot, y) : X \rightarrow E^N$ for a.a. $y \in Y$; in particular, $\tilde{X}(y) = \{x \in X : F(x, y) = 0\}$ is either ϕ or a differentiable manifold of dimension $(L - N)$ for a.a. $y \in Y$.

Remark 2.1. Let $\tilde{Y} = \{y \in Y : \tilde{X}(y) \text{ is either } \phi \text{ or a differentiable manifold of dimension } (L - N)\}$. Transversality theorem shows that $Y \setminus \tilde{Y}$ is a set of measure zero. $\tilde{Y}$ also enjoys an extra property. It

is *residual*, i.e., it is a countable intersection of open dense sets.

A useful corollary is the following.

Corollary 2.1. Let F be as in the transversality theorem and suppose $L < N$. Then $\bar{X}(y) = \phi$ for a.a $y \in Y$.

3 Linear Algebra and Control

Example 3.1. Let $m > n$, $A \in E^{m \times n}$, $b \in E^m$, and consider the solution of the over determined system, $Ax = b$ with b acting as a parameter. Let us set $L = n$, $X = E^n$, $M = N = m$, $Y = E^m$, and $F(x, b) = Ax - b$. Now, $JF = [A \mid -I_m]$, which has full row rank. Thus Corollary 2.1 implies that $Ax = b$ does not have a solution for a.a $b \in E^m$. Of course the result could have been directly obtained using the fact that $Ax = b$ has a solution iff b lies in the column space of A , a manifold whose dimension is at most n .

Example 3.2. We use Corollary 2.1 to show that $\det(A) \neq 0$ for a.a $A \in E^{n \times n}$. To set up A as a vector, let A^i denote the i -th column of A and $y = (A^1, A^2, \dots, A^n)$, a vector of length n^2 . We use the fact that $\det(A) = 0$ iff $\exists u \in E^n$ which satisfies $Au = 0$, $u^T u = 1$. Let us set $L = n$, $X = E^n$, $M = n^2$, $Y = E^{n^2}$, $N = n + 1$, and

$$F(u, y) = \begin{pmatrix} Au \\ (u^T u - 1)/2 \end{pmatrix}.$$

Let us study the Jacobian of $F(u, y)$. It is easy to check that

$$F_{(A^i, u_i)} = \begin{pmatrix} u_i I_n & A^i \\ 0 & u_i \end{pmatrix}.$$

Since $u^T u = 1$ implies that at least one of the u_i is non-zero, $JF(u, y)$ has full row rank $\forall (u, y) \in F^{-1}(0)$. By Corollary 2.1, for a.a $A \in E^{n \times n}$ the system $Au = 0$, $u^T u = 1$ does not have a solution, i.e., $\det(A) \neq 0$.

Example 3.3. We again use Corollary 2.1 to show that distinctness of the real roots of a polynomial is generic. Let $a = (a_1, a_2, \dots, a_n)$ and consider

$$p(s, a) = s^n + a_1 s^{n-1} + \dots + a_n.$$

Given a , $p(\cdot, a)$ will have a non-distinct real root iff $p(\cdot, a)$ and $p_s(\cdot, a)$ have a common real root. Let $L = 1$, $X = E$, $M = n$, $Y = E^n$, and $N = 2$. Define F by

$$\begin{aligned} F(s, a) &= \begin{pmatrix} p(s, a) \\ p_s(s, a) \end{pmatrix} \\ &= \begin{pmatrix} s^n + a_1 s^{n-1} + \dots + a_n \\ ns^{n-1} + (n-1)a_1 s^{n-2} + \dots + a_{n-1} \end{pmatrix}. \end{aligned}$$

Examining the last two columns of

$$F_a(s, a) = \begin{pmatrix} s^{n-1} & \dots & s & 1 \\ (n-1)s^{n-2} & \dots & 1 & 0 \end{pmatrix},$$

we see that $JF(s, a)$ always has rank two. By Corollary 2.1, $\{a : p(s, a) \text{ has a non-distinct real root}\}$ has measure zero in E^n .

achieve distinctness generically by perturbing only two of the coefficients, say (a_i, a_j) , and fixing all the others, say $a_k = \bar{a}_k \forall k \notin \{i, j\}$. In showing this, let us continue referring to the vector of coefficients formed from (a_i, a_j) and the $\bar{a}_k$ as a . There are several cases to be considered.

Case 1. $\{i, j\} = \{n, n-1\}$. Since the last two columns of F_a are always linearly independent it follows that

$$\{(a_i, a_j) : p(s, a) \text{ has a non-distinct real root}\} \quad (3.1)$$

has measure zero in E^2 .

Case 2. $\{i, j\} \cap \{n, n-1\}$ has a singleton, k . Take $k = n-1$ as an example. If $\bar{a}_n \neq 0$ then $F(0, a) \neq 0$, the columns of $F_a(s, a)$ corresponding to a_i and a_j are linearly independent $\forall (s, a) \in F^{-1}(0)$ and (3.1) follows. Even if $\bar{a}_n = 0$, (3.1) follows. To show this, we only have to reset $Y = \{(a_i, a_j) : a_{n-1} \neq 0\}$, an open set in E^2 , apply Corollary 2.1, and note that $\{(a_i, a_j) : a_{n-1} = 0\}$ is a set of measure zero in E^2 . If $k = n$, similar arguments hold. Thus (3.1) holds for case 2 also.

Case 3. $\{i, j\} \cap \{n, n-1\} = \phi$. If, either $\bar{a}_{n-1} \neq 0$ or $\bar{a}_n \neq 0$ then $F(0, a) \neq 0$, the columns of F_a corresponding to a_i and a_j are linearly independent, and (3.1) follows. If $\bar{a}_n = \bar{a}_{n-1} = 0$, then $s = 0$ is always a non-distinct root of $p(s, a)$. In that case we can take $X = E \setminus \{0\}$, an open set in E , apply Corollary 2.1 and obtain that $\{(a_i, a_j) : p(s, a) \text{ has a non-distinct real root which is non-zero}\}$ is a set of measure zero in E^2 .

Example 3.4. Consider the nonlinear autonomous system described by

$$\dot{x} = f(x) \quad (3.2)$$

where $f : E^n \rightarrow E^n$ and $f \in \mathcal{C}^2$. Let us consider a perturbed system,

$$\dot{x} = f(x) + Ax + b \quad (3.3)$$

where $A \in E^{n \times n}$ and $b \in E^n$ denote the parameters denoting the perturbation. Given (A, b) consider the system (3.3). A point, x^* is said to be an *equilibrium point* of (3.3) if $f(x^*) + Ax^* + b = 0$. An equilibrium point of (3.3), x^* , is said to be *hyperbolic* if $\det(f_x(x^*) + A) \neq 0$. We will use the transversality theorem and Corollary 2.1 to show that, for a.a (A, b) , each of the equilibrium points is isolated and hyperbolic. To set up A as a vector, let A^i denote the i -th column of A and $a = (A^1, A^2, \dots, A^n)$, a vector of length n^2 . Let Y be any open set in R^{n^2+n} . Y denotes the set of all perturbation vectors, (a, b) . If Y is a small set containing the origin, then one expects (3.2) and (3.3) to be close to each other. In this scenario the following result is useful.

Theorem 3.1. For a.a $(a, b) \in Y$, the following hold: (i) each equilibrium point of (3.3) is isolated; and (ii) each equilibrium point of (3.3) is hyperbolic.

Proof. We prove (i) using the transversality theorem. Set $L = n$, $X = E^n$, $M = n^2 + n$, Y as defined above, $N = n$, and $F(x, a, b) = f(x) + Ax + b$. Since $F_b = I_n$, it follows by the transversality theorem that

dimensional manifold. Note that a zero dimensional manifold is nothing but a set of isolated points. This proves (i).

We employ Corollary 2.1 to prove (ii). A point x is a non-hyperbolic equilibrium point of (3.3) iff $\exists u \in E^n$ such that the following are satisfied: $f(x) + Ax + b = 0$, $(f_x(x) + A)u = 0$, $u^T u - 1 = 0$. Let us set $L = 2n$, $X = E^{2n}$, $M = n^2 + n$, Y as in the statement of the theorem, $M = 2n + 1$ and

$$F(x, u, a, b) = \begin{pmatrix} f(x) + Ax + b \\ (f_x(x) + A)u \\ (u^T u - 1)/2 \end{pmatrix}.$$

Since $u^T u = 1$, at least one component of u is non-zero, say u_i . Now

$$F_{(b, A^i, u_i)}(x, u, a, b) = \begin{pmatrix} I_n & x_i I_n & 0 \\ 0 & u_i I_n & A^i \\ 0 & 0 & u_i \end{pmatrix},$$

a non-singular matrix. Thus $JF(x, u, a, b)$ has full row rank $\forall (x, u, a, b) \in F^{-1}(0)$. By Corollary 2.1, for a.a. $(a, b) \in Y$, $F = 0$ does not have a solution and so (ii) is also proved. ■

4 Optimization

Homotopy methods. Homotopy is a powerful tool for solving a square system of nonlinear equations [2, 10, 12]. A well-known application of the transversality theorem is in showing that the zero manifold associated with a homotopy method is generically regular. Here we will briefly illustrate the ideas for the fixed point homotopy [2, 12]. Let $f : E^n \rightarrow E^n$ be in C^2 and consider the solution of $f(x) = 0$. Introduce a homotopy variable, $\lambda \in E$ and a parameter, $a \in Y$, an open set in E^n to get the homotopy function, $F(x, \lambda, a) = \lambda f(x) + (1 - \lambda)(x - a)$. Given $a \in Y$, let $W(a) = \{(x, \lambda) : F(x, \lambda, a) = 0\}$ denote the zero manifold of $F(\cdot, \cdot, a)$. The homotopy method consists of choosing an $a \in Y$ and tracking $W(a)$ from the known point $x = a$, $\lambda = 0$, until the hyperplane $\lambda = 1$ is hit. The value of x when this occurs is a solution of $f(x) = 0$.

An important requirement for the homotopy method to work well is that $W(a)$ is a one dimensional manifold. The transversality theorem is useful here. Let $L = n + 1$, $X = R^n \times (-\infty, 1)$, $M = n$, Y as mentioned earlier, and $N = n$. Since $F_a = -(1 - \lambda)I_n$, $JF(x, \lambda, a)$ has full row rank $\forall (x, \lambda, a) \in X \times Y$. Further, $W(a) \neq \emptyset$ since it contains $(a, 0)$. Thus by the transversality theorem, $W(a)$ is a one dimensional manifold for a.a. $a \in Y$.

Nonlinear Programming. Consider the problem,

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && h(x) = 0, \quad g(x) \leq 0, \end{aligned} \quad (4.1)$$

where $x \in E^n$, $f : E^n \rightarrow E$, $h : E^n \rightarrow E^m$, $g : E^n \rightarrow E^l$, and, $f, g, h \in C^2$. Consider the following perturbation of (4.1):

$$\begin{aligned} &\text{minimize} && f(x) + \alpha^T x + x^T Bx \\ &\text{subject to} && h(x) + \gamma = 0, \quad g(x) + \delta \leq 0, \end{aligned} \quad (4.2)$$

are perturbation parameters. Let $\beta \in E^n$ denote the vector containing the elements of B . Then, $y = (\alpha, \beta, \gamma, \delta) \in R^{n+n^2+m+l}$ denotes the vector of perturbation parameters. Let us choose small $\bar{y}_i > 0$, define

$$Y = \{y : -\bar{y}_i < y_i < \bar{y}_i \forall i\}, \quad (4.3)$$

and restrict y to the set Y so that (4.1) and (4.2) will be "close" to each other for any $y \in Y$. The question, 'how close are (4.1) and (4.2)?' is an issue of the stability of (4.1) which will not be discussed here. See [4] for results on stability. Our main aim here is to study the generic relationship between necessary and sufficient conditions of optimality using the transversality theorem.

For $y \in Y$, define

$$\mathcal{F}(y) = \{x : h(x) + \gamma = 0, \quad g(x) + \delta \leq 0\}. \quad (4.4)$$

If there exists an $\tilde{x}$ satisfying $h(\tilde{x}) = 0$, $h_x(\tilde{x})$ has full row rank, and $g(\tilde{x}) < 0$, then it is easy to see that, if each $\bar{y}_i$ is chosen to be sufficiently small (but still, positive) then $\mathcal{F}(y) \neq \emptyset \forall y \in Y$, so that (4.2) is not a null problem.

Let y be fixed, $x \in \mathcal{F}(y)$ and $I(x) = \{j : g_j(x) + \delta_j = 0\}$, the index set of active inequality constraints at x . We say x is a *regular point* of $\mathcal{F}(y)$ if the gradient vectors of h_i , g_j , $1 \leq i \leq m$, $j \in I(x)$ at x are linearly independent. $\mathcal{F}(y)$ is said to be *regular* if every $x \in \mathcal{F}(y)$ is a regular point. Define the Lagrangian function, L as

$$L(x, \lambda, \mu; y) = f(x) + \alpha^T x + x^T Bx + \lambda^T h(x) + \mu^T g(x).$$

The following theorems condense the well-known necessary and sufficient conditions of optimality, due to Karush, Kuhn and Tucker [7].

KKT Theorem (NC). Let $y \in Y$ be fixed and suppose x is a local minimum of (4.2) as well as a regular point of $\mathcal{F}(y)$. Then $\exists \lambda \in E^m$ and $\mu \in E^l$ such that the following hold: (i) $L_x(x, \lambda, \mu; y) = 0$; (ii) $\mu^T g(x) = 0$; (iii) $\mu \geq 0$; and (iv) $L_{xx}(x, \lambda, \mu; y)$ is positive semidefinite on $T(x)$, the tangent space of active constraints at x , defined by

$$T(x) = \{z : h_x(x)z = 0, \quad g_{j_x}(x)z = 0 \forall j \in I(x)\}. \quad (4.5)$$

KKT Theorem (SC). Let $y \in Y$ be fixed and, $x \in \mathcal{F}(y)$ is such that $\exists \lambda \in E^m$, $\mu \in E^l$ and the following conditions are satisfied: (i) $L_x(x, \lambda, \mu; y) = 0$; (ii) $\mu^T g(x) = 0$; (iii) $\mu_j > 0 \forall j \in I(x)$; and (iv) $L_{xx}(x, \lambda, \mu; y)$ is positive definite on $T(x)$, where $T(x)$ is as in (4.5). Then x is a local minimum of (4.2).

A point $x \in \mathcal{F}(y)$ which satisfies conditions (i)-(iii) of KKT Theorem (NC) will be called as a *KKT point*. A corresponding λ and μ will be referred to as *Lagrange multipliers* associated with x . Following is the main result that we will prove using the transversality theorem.

Theorem 4.1. For a.a. $y \in Y$, $x \in \mathcal{F}(y)$ is a local minimum of (4.2) iff it satisfies conditions (i)-(iv) of KKT Theorem (NC).

The proof of Theorem 4.1 is based on the following lemma.

Lemma 4.2. For a.a. $y \in Y$ the following properties hold:

- (b) if x is a KKT point of (4.2) and, λ, μ are Lagrange multipliers associated with x , then $\lambda_i \neq 0 \forall 1 \leq i \leq m$, and $\mu_j \neq 0 \forall j \in I(x)$; and,
- (c) for every KKT point, x and any Lagrange multipliers λ and μ associated with it, $L_{xx}(x, \lambda, \mu; y)$ is non-singular.

Proof. Let us say that $x, z \in \mathcal{F}(y)$ belong to the same active set type if $\{j : g_j(x) + \delta_j = 0\} = \{j : g_j(z) + \delta_j = 0\}$, i.e., the index set of active inequalities is the same for x and z . Since I , the number of inequalities is finite, the number of active set types is also finite. Because of this and the fact that the kind of perturbations for equalities and inequalities is similar (i.e., each equality or inequality has a separate perturbation variable γ_i or δ_j associated with it) it is sufficient if we prove the lemma for the problem with equalities only. (Note that the arguments for the 'equalities only' case can be repeated for each active set type and the resulting conclusions can be combined using Lemma 1.1.) Thus, for the rest of the proof given below, we assume no inequalities are present and take $y = (\alpha, \beta, \gamma)$.

(a) Let $L = X = E^n$, $M = n + n^2 + m$, Y as in the statement of the lemma, i.e., as in (4.3), $N = m$, and define F by $F(x, y) = h(x) + \gamma$. Since $F_\gamma = I_m$, JF always has full row rank. By the transversality theorem $\mathcal{F}(y)$ is regular for a.a $y \in Y$.

(b) Take any i such that $1 \leq i \leq m$, and look at the possibility of λ_i becoming zero. Let $L = n + m$, $X = E^{n+m}$, M and Y as in the proof of part (a) given above, $N = n + m + 1$, and define F by

$$F(x, \lambda, y) = \begin{pmatrix} L_x^T = f_x^T + \alpha + (B + B^T)x + h_x^T \lambda \\ h(x) + \gamma \\ \lambda_i \end{pmatrix}.$$

Now, $F_{(\alpha, \gamma, \lambda_i)}$ is an upper triangular matrix with all diagonal elements equal to 1. Thus JF always has full row rank. By Corollary 2.1, $\lambda_i \neq 0$ for a.a $y \in Y$. Combining this result for all i and using part (a) of Lemma 1.1 we obtain part (b) of Lemma 4.2.

(c) Here we use the fact that L_{xx} is singular iff $\exists u \in E^n$ such that $L_{xx}u = 0$, $u^T u - 1 = 0$. Let $L = n + m + n$, $X = E^{n+m+n}$, M and Y as mentioned earlier, $N = n + m + n + 1$, and define F by

$$F(x, \lambda, y) = \begin{pmatrix} L_x^T = f_x^T + \alpha + (B + B^T)x + h_x^T \lambda \\ h(x) + \gamma \\ L_{xx}u \\ (u^T u - 1)/2 \end{pmatrix}.$$

Consider any (x, λ, u, y) such that $F(x, \lambda, u, y) = 0$. Since $u^T u - 1 = 0$, at least one component of u is non-zero, say $u_j \neq 0$. Let B^j = the j -th column of B , which is a part of β , which in turn is a part of y . At (x, λ, u, y) , the sub-Jacobian, $F_{(\alpha, \gamma, B^j, u_j)}$ has the following structure:

$$F_{(\alpha, \gamma, B^j, u_j)} = \begin{pmatrix} I_n & 0 & \times & 0 \\ 0 & I_m & 0 & 0 \\ 0 & 0 & T & \times \\ 0 & 0 & 0 & u_j \end{pmatrix}.$$

portant for our analysis here, $T = (L_{xx}u)_{B^j} = u_j I_n + eu^T$, and e = unit vector in E^n with $e_i = 0$, $i \neq j$ and $e_j = 1$. T is non-singular since $\det(T) = 2(u_j)^n \neq 0$. Thus $F(\alpha, \gamma, B^j, u_j)$ is non-singular. Hence JF has full row rank $\forall (x, \lambda, u, y) \in F^{-1}(0)$. By Corollary 2.1 we have (c). ■

Proof of Theorem 4.1. Let $\tilde{Y} = \{y \in Y : \text{properties (a)-(c) of Lemma 4.2 hold}\}$. By Lemma 4.2 and part (a) of Lemma 1.1, $Y \setminus \tilde{Y}$ has zero measure. Let $y \in \tilde{Y}$ and consider (4.2). By part (a) of Lemma 4.2, conditions (i)-(iv) of KKT Theorem (NC) are necessary for a local minimum. Let us show that they are also sufficient. Part (b) of Lemma 4.2 and condition (iii) of KKT Theorem (NC) imply condition (iii) of KKT Theorem (SC). Also, part (c) of Lemma 4.2 and condition (iv) of KKT Theorem (NC) imply condition (iv) of KKT Theorem (SC). Thus by KKT Theorem (SC), conditions (i)-(iv) of KKT Theorem (NC) are also sufficient for a local minimum. ■

Degeneracy in Linear Programming. We will show that the absence of degenerate basic feasible solutions in linear programs is, in some sense, generic. Let $A \in E^{m \times n}$ have full row rank and consider

$$\mathcal{F}(b) = \{x \in E^n : Ax = b, x \geq 0\},$$

the feasible set of a linear program in standard form looked at as a function of $b \in E^m$. A point x is said to be a *basic feasible solution* (BFS) of $\mathcal{F}(b)$ if: (i) $x \in \mathcal{F}(b)$; (ii) at least $(n - m)$ components of x are zero; and (iii) the columns of A corresponding to the non-zero elements of x are linearly independent. A BFS, x is said to be *degenerate* if more than $(n - m)$ of its components are zero.

Theorem 4.3. $\mathcal{F}(b)$ does not have any degenerate BFS for a.a $b \in E^m$.

Proof. Let $I \subset \{1, \dots, n\}$ be any index set of cardinality $(m - 1)$ such that the columns of A corresponding to the indices in I are linearly independent. Let $\tilde{A} \in E^{m \times (m-1)}$ be the matrix whose columns are the columns of A corresponding to the indices in I . By the result of Example 3.1 it follows that $\tilde{A}\tilde{x} = b$ does not have a solution for a.a $b \in E^m$. By part (a) of Lemma 1.1 and the fact that the number of index sets of the type I is finite, the theorem follows. ■

5 Geometry

In the field of Computational Geometry, it is a common practice to make generic assumptions which help avoid troublesome situations that require special case by case analysis. In the literature these assumptions are usually called as general position assumptions. In most papers the genericity of such assumptions is never established because it is intuitively obvious. In this section we take up three general position assumptions and establish their genericity using the transversality theorem.

Example 5.1. In the construction of the Voronoi diagram of points in E^n [9, 3] the following general position assumption is usually made: $\{x^1, \dots, x^k\}$ ($k \geq n + 2$), a given set of distinct data points in E^n have the property that there does not exist a sphere which contains $(n + 2)$ or more of the data points on it. We will use Corollary 2.1 to show

$x \in E$, $i = 1, \dots, k$. It is sufficient if we demonstrate the above for $k = n + 2$ only because: the number of subsets of $\{x^1, \dots, x^k\}$ with cardinality equal to $(n + 2)$ is finite; the ideas for $\{x^1, \dots, x^{n+2}\}$ can be repeated for each of these subsets; and, the genericity result for each of the subsets can be combined using Lemma 1.1 to yield the genericity result for $\{x^1, \dots, x^k\}$.

Let us consider the existence of a sphere, $S = \{x : (x - c)^T(x - c) = r^2\}$ which contains the x^i , $i = 1, \dots, n + 2$, on it. Let $L = n + 1$, $X = E^{n+1}$, $M = n(n + 2)$, $y = (x^1, \dots, x^{n+2})$, $Y = \{y : y = (x^1, \dots, x^{n+2}) \text{ and the } x^i \text{ are distinct}\}$, and $N = n + 2$. Note that Y is open in E^M . Define F as follows. For $i = 1, \dots, n + 2$, define the i -th component of F , F_i as $F_i(c, r, y) = [(x^i - c)^T(x^i - c) - r^2]/2$. Let $(c, r, y) \in F^{-1}(0)$. Since the x^i are distinct, $r > 0$. This means that $x^i - c \neq 0 \forall i$. $F_y(c, r, y)$ is a block diagonal matrix with $(x^i - c)^T$, $i = 1, \dots, n + 2$, as the diagonal blocks, and so it has full row rank. Our result then follows by Corollary 2.1.

Example 5.2. Consider $(n + 2)$ surfaces, $\{C_i\}$ in E^n represented by

$$C_i = \{x \in E^n : h_i(x) = 0\}$$

where $h_i : E^n \rightarrow E$ and $h_i \in C^2$, $i = 1, \dots, n + 2$. We will show that generically there will not exist a sphere with positive radius which will touch all the $(n + 2)$ surfaces. This property is useful in the construction of generalized Voronoi diagrams of nonlinear objects. To establish this property we consider the following perturbation of h_i for each $i = 1, \dots, n + 2$:

$$\tilde{h}_i(x, p^i, q^i) = h_i(x) + p^{iT}x + q^i.$$

Let $y = (p^1, q^1, p^2, q^2, \dots, p^{n+2}, q^{n+2}) \in E^{(n+1)(n+2)}$, the vector of all perturbation variables, belong to some open set, Y . If Y is some small open set containing the origin then for each i , the perturbed surface $\tilde{C}_i$ defined by

$$\tilde{C}_i = \{x \in E^n : \tilde{h}_i(x, p^i, q^i) = 0\}$$

will be "close" to C_i for any given $y \in Y$.

Let $S = \{x : (x - c)^T(x - c) - r^2 = 0\}$ denote a sphere which touches the surfaces $\tilde{C}_1, \dots, \tilde{C}_{n+2}$ at $x^1, \dots, x^{n+2}$, respectively. S touches $\tilde{C}_i$ at x^i iff $\exists \alpha^i \in E$ such that

$$f^i(x^i, \alpha^i, c, r, p^i, q^i) = 0,$$

where

$$f^i = \begin{pmatrix} \tilde{h}_i \\ \alpha^i(x^i - c) - \nabla_{x^i} \tilde{h}_i \\ [(x^i - c)^T(x^i - c) - r^2]/2 \end{pmatrix},$$

and $\nabla_{x^i} \tilde{h}_i$ is the gradient of $\tilde{h}_i$ with respect to x^i . Let $L = (n + 1)(n + 3)$, $x = (x^1, \alpha^1, \dots, x^{n+2}, \alpha^{n+2}, c, r)$,

earlier, $N = (n + 2)(n + 2)$ and F defined by

$$F(x, y) = \begin{pmatrix} f^1 \\ \vdots \\ f^{n+2} \end{pmatrix}$$

Consider A^i , the Jacobian of f^i with respect to $w^i = (q^i, p^i, x^i)$. A^i has the following structure:

$$A^i = \begin{pmatrix} 1 & \times & \times \\ 0 & -I_n & \times \\ 0 & 0 & (x^i - c)^T \end{pmatrix}$$

where the $\times$ denotes matrix functions which are unimportant. Since $r > 0$, $x^i - c$ is a non-zero vector. Thus A^i has full row rank $\forall (x, y) \in F^{-1}(0)$. Furthermore, if we set $w = (w^1, \dots, w^{n+2})$, then A , the Jacobian of F with respect to w , is a block diagonal matrix with $A^1, \dots, A^{n+2}$ as the diagonal blocks. Because each A^i has full row rank, A also has full row rank. Thus, $JF(x, y)$ has full row rank $\forall (x, y) \in F^{-1}(0)$. By Corollary 2.1 $F(x, y) = 0$ does not have a solution for a.a $y \in Y$ and our result that for a.a $y \in Y$ there does not exist a sphere with positive radius which touches each of $\tilde{C}_i$, $i = 1, \dots, n + 2$, follows.

If $r = 0$ is allowed then it corresponds to the condition that all the $\tilde{C}_i$ have a common point. By a slight change of the earlier arguments it can be shown that generically this also cannot occur. In fact it can be shown that, generically even $(n + 1)$ surfaces will not intersect.

Example 5.3. We consider a general position assumption used in robot motion planning [8]. Suppose $P^1, \dots, P^k$ are k non-intersecting convex polygons in E^2 representing obstacles, and P^0 be a convex polygon which denotes the moving object. For $i = 0, 1, \dots, k$, let: n^i = the number of vertices of P^i ; and P^i be represented by $\{v^{i,j} \in E^2 : j = 1, \dots, n^i\}$, the set of vertices of P^i given in clockwise order. A convex polygon, $\tilde{P}^0$ is said to be a *homothetic copy* of P^0 if $\exists z \in E^2$ such that $\tilde{P}^0 = P^0 + \{z\}$, where $+$ denotes the Cartheodory-Minkowski set sum. Two convex polygons are said to be *touching* if they intersect but their interiors do not.

Leven and Sharir [8] have given an algorithm for moving P^0 from any initial position to any final position by translation without colliding with any of the obstacle polygons. To simplify the description of their algorithm they make the following general position assumption: there does not exist a homothetic copy of P^0 which touches more than two members of $\{P^1, \dots, P^k\}$. Here we will show that the above assumption is generic.

Let us first give a precise statement of our result. Let $y^i = (v^{i,1}, \dots, v^{i,n^i})$, $i = 0, 1, \dots, k$, $M = 2 \sum_{i=0}^k n^i$, $y = (y^0, y^1, \dots, y^k)$, and

$$Y = \{y \in E^M : \begin{array}{l} P^i \text{ is a convex polygon with } n^i \\ \text{vertices } \forall i = 0, 1, \dots, k, \text{ and} \\ P^i \cap P^j = \emptyset \forall \{i, j\} \subset \{1, \dots, k\}, \\ i \neq j. \end{array}\}$$

this note the following: the condition, ' P is a convex polygon with n^i vertices' corresponds to the requirement that the vertices of P^i are distinct and that, for each three consecutive vertices of P^i , say, $\{a, b, c\}$, b lies strictly to the left of the directed line from a to c ; the condition, $P^i \cap P^j = \emptyset$ corresponds to the requirement that $d(P^i, P^j) = \min \{\|\alpha - \beta\| : \alpha \in P^i, \beta \in P^j\} > 0$; and, each of these requirements can be written as a strict inequality of the form $\rho(y) > 0$ or $\rho(y) \neq 0$ where ρ is a continuous function of y . Our main result here is that, for a.a $y \in Y$ there does not exist a homothetic copy of P^0 which touches more than two members of $\{P^1, \dots, P^k\}$.

Let $\tilde{P}^0 = P^0 + \{z\}$ be a homothetic copy of P^0 . If $\tilde{P}^0$ touches some P^i , $i \neq 0$, it will involve a vertex of one of $\{P^0, P^i\}$ lying on an edge (line segment joining two consecutive vertices) of the other. Let us consider the occurrence of three such touches. There are several ways in which the three touches can occur. We will take one difficult case and demonstrate below, genericity of the non-occurrence of it. Repeating the ideas for each of the cases, which are finite in number, and using Lemma 1.1 then yields the overall result that we wish to prove.

The case that we consider is the following. Let $\{a, b, c\}$ be three consecutive vertices of P^0 such that: a vertex, d of P^1 lies on the edge joining $a + z$, $b + z$; a vertex, e of P^2 lies on the edge joining $b + z$, $c + z$; and, $b + z$ lies on the edge joining two consecutive vertices, f and g of P^3 . Note that (a, b, c, d, e, f, g) is a subvector of y , the vector of data for the motion planning problem. Introduce three real variables α , β and γ , set $L = 5$, $x = (z, \alpha, \beta, \gamma)$, $X = E^5$, $N = 6$, define F by $F(x, y) = (d - \tilde{a} - \alpha(\tilde{b} - \tilde{a}), e - \tilde{b} - \beta(\tilde{c} - \tilde{b}), \tilde{b} - f - \gamma(g - f))$, where $(\tilde{a}, \tilde{b}, \tilde{c}) = (a + z, b + z, c + z)$, and note that if, for any given y , the three touches occur then $F(x, y) = 0$ has a solution, x . Thus if we show that $F(x, y) = 0$ does not have a solution for a.a $y \in Y$, then the case we consider (i.e., $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $0 \leq \gamma \leq 1$) also cannot occur for a.a $y \in Y$. It is easy to see that $F_{(d, e, b)}$ is an upper triangular matrix all of whose diagonal elements are unity. Thus JF has full row rank and, by Corollary 2.1, $F(x, y) = 0$ has no solution for a.a $y \in Y$.

6 Conclusion

We have demonstrated, by way of examples in linear algebra and control, optimization, and geometry, that the transversality theorem of differential topology is a very useful tool for establishing the genericity of a property which is dependent on a finite number of real parameters and which is expressible using a system of nonlinear equations. Some of the results derived in this paper (e.g., Example 5.2) are non-trivial to establish by direct means, without resorting to the use of the transversality theorem.

References

- [1] R. A. Abraham and J. Robbin, *Transversal Mappings and Flows*, Benjamin, New York, 1976.

"Finding Zeroes of Maps: Homotopy Methods that are Constructive with Probability one", Math. Comput., vol. 32, pp. 887-899, 1978.

- [3] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, Springer Verlag, 1987.
- [4] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Non-linear Programming*, Academic Press, New York, 1983.
- [5] V. Guillemin and A. Pollack, *Differential Topology*, Prentice Hall Inc., New Jersey, 1974.
- [6] M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press, New York, 1974.
- [7] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Addison-Wesley, Reading, Mass, 1984.
- [8] D. Leven and M. Sharir, "Planning a Purely Translational Motion for a Convex Object in Two Dimensional Space Using Generalized Voronoi Diagrams", Discrete and Computational Geometry, vol. 2, pp 9-31, 1987.
- [9] F. P. Preparata and M. I. Shamos, *Computational Geometry - An Introduction*, Springer Verlag, New York, 1985.
- [10] S. L. Richter and R. A. De Carlo, "Continuation Methods: Theory and Applications", IEEE Trans. Auto. Control, vol. AC-28, pp 660-665, 1983.
- [11] H. L. Royden, *Real Analysis*, The McMillan Co., New York, 1963.
- [12] L. T. Watson, *A globally convergent algorithm for computing fixed points of C^2 maps*, Appl. Math. Comput., Vol. 5, 1979, pp 297-311.