Original Research Paper

A Geometric Generalization of the Planar Gale-Nikaidô Theorem

E. Cabral Balreira

Department of Mathematics, Trinity University, San Antonio, Texas, USA

Abstract: The Gale-Nikaidô Theorem establishes global injectivity of maps defined over rectangular regions provided the Jacobian matrix is a *P*-matrix. We provide a purely geometric generalization of this result in the plane by showing that if the image of each edge of the rectangular domain is realized as a graph of a function over the appropriate axis, then the map is injective. We also show that the hypothesis that the Jacobian matrix is a *P*-matrix is simply one way to analytically check this geometric condition.

Keywords: Global Injectivity, *P*-Matrix, Gale-Nikaidô

Introduction

The study of global injectivity results has important applications in Mathematical Economics, in particular problems in International Trade, see Samuelson (Samuelson, 1949; 1967). One of the main questions is to determine the conditions on the system of cost functions, denoted by F, so that the equation $F(x) = y$ has at most one solution, where x is the input factor price and y represents cost, see the work of Mas-Colell (1979) and Parthasarathy (1983) for more details.

In this study, we are motived with the Gale and Nikaidô (1965) global univalence theorem. It states that for a rectangular region given by $R = \{x \mid a_i \le x_i \le b_i, i = 1,..., n\}$ for $a_i < b_i$ and $F: R \to \mathbb{R}^n$ an orientation preserving local diffeomorphism for which all principal minors of the Jacobian matrix are everywhere positive, then F is injective. Our focus is to provide a purely geometric generalization for the planar case. As highlighted by Aleksandrov (1994), the Gale-Nikaidô Theorem has yet to be given a geometric proof, or even an interpretation. However, recently an elementary approach has been considered by (Fujimoto and Herrero, 2000). In addition, it was also shown by Aleksandrov (1994) that injectivity of mappings defined in the rectangular domains cannot be extended to mappings defined in arbitrary convex domains. We expect that our result can provide better insight for the Gale-Nikaidô Theorem in higher dimensions.

Let $R = [0, 1]^2 \subseteq \mathbb{R}^2$ be the unit square in the plane. Consider a decomposition of the boundary of R, denoted by ∂R, where the 4 edges are labeled $R_j^{\delta_j}$ for $j = 1, 2$ and $\delta_j = 0$ or 1 and $R_j^{\delta_j} = \{(x_1, x_2) \in R \mid x_j = \delta_j\}$.

Given $F: R \to \mathbb{R}^2$ a map of class C^1, we say that F *normally assembles R* if, up to relabeling, for each for each edge $R_j^{\delta_j}$ of R, we have that its image $F\left(R_j^{\delta_j}\right) = \varphi_j^{\delta_j}$ is a graph over the x_i-axis, for $i \ne j$ where the unit normal at an interior point of $\varphi_j^{\delta_j}$ has the same direction in the jth component as in the domain. More precisely, let α be a smooth curve in the plane and denote the normal to γ at P to be $\eta(P; \gamma)$, then F normally assembles R if for all $P \in \text{int}\left(\varphi_j^{\delta_j}\right)$, we have:

$$\left\langle \eta\left(P; \varphi_j^{\delta_j}\right), \left(1 - 2\delta_j\right)e_j \right\rangle > 0 \qquad (*)$$

where, $\{e_1, e_2\}$ is the canonical orthonormal basis of $\mathbb{R}^2$. We also must consider a relabeling of the edges to account for a change of basis, for instance, the map F: $[0, 1] \times [0, 1] \to \mathbb{R}^2$ given by $F(x, y) = (-x, y)$ normally assembles the unit square after we relabel R_2^0 with R_2^1 and vice-versa.

Our main result is the following.

Theorem 1.1

Let R be the unit square and $F: R \to \mathbb{R}^2$ be a C^1 local diffeomorphism. If F normally assembles R, then F is injective.

The geometric condition in the hypotheses of Theorem 1.1 can be analytically checked, as it will be shown in Proposition 2.4. Thus we will have shown that the Gale-Nikaidô Theorem (Gale and Nikaidô, 1965) is a

corollary of Theorem 1.1 which now has a purely geometric proof.

Proof of Geometric Planar Results

We will now provide the proof of our main result in Theorem 1.1. We will show that F will be in injective on the boundary of R which is enough to show that F is injective on the entire domain by a result of Kestelman (1971). Let us recall.

Theorem 2.1 (Kestelman (1971))

Let $F: K \to \mathbb{R}^n$ be an open and locally injective map. If $K \subseteq \mathbb{R}^n$ is a compact set, ∂K is connected and $F|_{\partial K}$ is injective, then F is injective.

In addition, we will introduce a concept of exposed points, first introduced by the by Balreira *et al.* (2014), to determine regions of injectivity of discrete dynamical systems.

Definition 2.2

Let $U \subseteq \mathbb{R}^2$ be a compact region, $p \in U$ and $v \in S^1$, that is, a point in the unit circle. We say that p is exposed in the direction of v if there exists $\varepsilon > 0$ such that the ray $r_v(t) = p + tv \in U$ for $t \in (0, \varepsilon)$.

From the definition, we can see that if $p \in int(U)$, then p is exposed in every direction. In our applications, the notion of exposed points will be used to show that the ray in the direction v eventually will hit the boundary of U. Namely, if $p \in U$ is exposed in the direction of v, then for some $t > 0$, $r_v(t) \in \partial U$.

We will only be looking at the directions $\pm e_1$ and $\pm e_2$. Then in order to check if $P \in \gamma$ is exposed in the direction of δe_j is equivalent to check, for $\delta = 0$ or 1, if:

$$\left\langle \eta(P; \gamma), (1 - 2\delta)e_j \right\rangle > 0 \qquad (2.1)$$

This amounts to looking at the sign of the jth coordinate of $\eta(P; \gamma)$, which can be done by checking if the following holds:

$$\text{sgn}\left\{ \pi_j \left(\eta(P; \gamma) \right) \right\} = (1 - 2\delta) \qquad (2.2)$$

where, π_j is the projection onto the coordinate and sgn is the sign function such that $\text{sgn}(t) = +1$, if t 0, $\text{sgn}(0) = 0$ and $\text{sgn}(t) = -1$, if $t < 0$.

Let us now consider a result indicating that the graph of a function only changes orientation as it crosses a critical point. Although this is a straightforward result, it sets the notation that will be used in the main result.

Lemma 2.3

Let $f: [a, b] \to \mathbb{R}$ be a C^1 map. Suppose $z \in \mathbb{R}$ is a regular value of f such that $f(a) \neq z$. Consider $\Gamma_f = \{(x,$ $f(x))| \; x \in [a, b]\}$, the graph of the map f and $c = \min\{f^{-1}(z)\} \neq \emptyset$. Then for $P = (c, f(c))$:

$$\text{sgn}\left\{ \pi_1 \left(\eta(P; \Gamma_f) \right) \right\} = \text{sgn}\left\{ f(a) - z \right\} or\left(\Gamma_f \right)$$

where, $or(\gamma) = +1$ if $\{(1, f'(c)), \eta(P; \Gamma_f)\}$ is a positively oriented basis of $\mathbb{R}^2$ and $or(\Gamma_f) = -1$ otherwise.

Proof

We may assume that the orientation of $\gamma \subset \mathbb{R}^2$ is the canonical one, that is, the normal at each point is a positive $\frac{\pi}{2}$ degree rotation of the tangent to γ. For simplicity, let the graph of f, denoted by Γ_f be given by the curve $\gamma: [a, b] \to \mathbb{R}^2$ where $\gamma(t) = (t, f(t))$. Thus for $t \in [a, b]$, we have $\eta((t, f(t)); \gamma) = (-f'(t), 1)$. In particular, $\pi_1(\eta(P; \gamma)) = -f'(c)$.

Without loss of generality, let us assume that $z > f(a)$. The case where $z < f(a)$ may be treated in analogous manner by a reflection of the graph of f about the x_1-axis.

We claim that for $c = \min\left\{f^{-1}(z)\right\} \neq \emptyset$, we must have $f'(c) > 0$. Indeed, since z is a regular value $f'(c) \neq 0$. So if it is not the case that $f'(c) > 0$, we must have that $f'(c) < 0$ and f is locally decreasing at c. Therefore, there is d such that $a < d < c$ with $f(d) > z$ as depicted in Fig. 1.

Then by the Intermediate Value Theorem to f restricted to $[a, d]$, we find $c' \in (a, d)$ with $f(c') = z$, a contradiction to the definition of c as the smallest pre-image of z. Thus, $f'(c) > 0$ and:

$$\begin{aligned} &\text{sgn}\left\{ \pi_x \left(\eta(P; \gamma) \right) \right\} = \text{sgn}\left\{ -f'(c) \right\} = \\ &-1 = (-1)(+1) = \text{sgn}\left\{ f(a) - z \right\} or\left(\lambda \right) \end{aligned} \qquad (2.3)$$
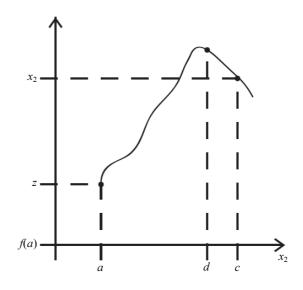

Fig. 1: Behavior of the graph of a function before the first critical point

We are now ready to give the proof of the main result.

Proof of Theorem 1.1

Let R_1^0, R_1^1, R_2^0 and R_2^1 be the cubical assemblage of R and $F\left(R_j^{\delta_j}\right)=\varphi_j^{\delta_j}$.

We will show that $F|_{\partial\Omega}$ is injective, then by (Kestelman, 1971) it will follow that F is injective. The main geometric idea is to show that whenever the interior of two edges intersect, at the intersection there will be an exposed point. Thus if we follow the ray at such exposed point, at every time it intersects the boundary again, it will continue to create exposed points as long as $\varphi_j^{\delta_j}$ can be viewed as graphs of functions with the correct orientation.

Without loss of generality, it suffices to consider just one edge, say φ_2^0 and all others can be done similarly. Assume that φ_2^0 as a graph over the x_1-axis with endpoints a_0 and a_1. Let $u = (u_1, u_2)\in \varphi_2^0$ be an interior point and consider the vertical ray in the direction of $-e_2$ starting at u denoted by r_u: $[0,\infty) \to \mathbb{R}^2$ given by $r_u(t) = u - te_2 = (u_1, u_2 - t)$.

We will show that for $t \geq 0$, $r_u(t) \notin \left(\varphi_2^1\cup\varphi_1^0\cup\varphi_1^1\right)$. Indeed, suppose towards a contradiction that this is not the case. Then at least one of the following possibilities must hold. The ray must intersect the edge above, or one at each side. Let us consider each scenario separately.

Suppose for some $t\geq 0$, $r_u(t)\in \varphi_2^1$. Since φ_2^1 is a graph over the x_1-axis, there is a unique $t_2^1\geq 0$ so that $r_u\left(t_2^1\right)\in\varphi_2^1$. Let $P_2^1 = r_u\left(t_2^1\right)=\left(u_1,w_2^1\right)$ as depicted in Fig. 2. Since F normally assembles R, we have:

$$\left\langle\eta\left(P_2^1;\varphi_2^1\right),(1-2\cdot1)e_2\right\rangle>0 \qquad (2.4)$$

Thus:

$$\mathrm{sgn}\left\{\pi_2\left(n\left(P_2^1;\varphi_2^1\right)\right)\right\}=-1 \qquad (2.5)$$

Hence the point P_2^1 is exposed in the direction $-e_2$ as depicted in Fig. 2. Therefore, the ray r_u must intersect $\partial F(R)$ at a point $r_u(t)$ with $t>t_2^1$. If this case does not happen, for notation purposes, let us set $t_2^1=0$.

Suppose there is $t\geq0$ such that $r_u(t)\in\varphi_1^0$. Let $t_1^0=\max\left\{t\right|r_u\left(t\right)\in\varphi_1^0\right\}$ and denote the point $P_1^0 = r_u\left(t_1^0\right)=\left(u_1,w_1^0\right)$. As we consider φ_1^0 as a graph of a function over the x_2-axis, let us denote this map by $g_1^0:[a,b]\to\mathbb{R}$, with $g_1^0\left(a\right)=a_0$ as depicted in Fig. 3.
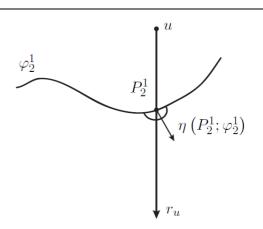

Fig. 2: Intersection between ray $r_u(t)$ and φ_2^1 creating an exposed point P_2^1
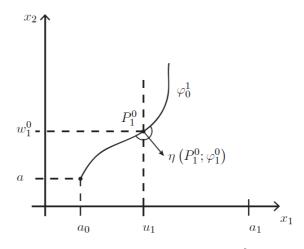

Fig. 3: Intersection between ray $r_u(t)$ and φ_1^0 creating an exposed point P_1^0

First notice that $w_1^0\in(a,b)$ and from the choice of t_1^0, we have that $w_1^0 = \min\left\{\left(g_1^0\right)^{-1}\left(u_1\right)\right\}$.

Next, observe that the orientation of φ_1^0 is the canonical positive orientation as F normally assembles R. Then, an application of Lemma 2.3 yields the following:

$$\mathrm{sgn}\left\{\pi_2\left(\eta\left(P_1^0;\varphi_1^0\right)\right)\right\}=\mathrm{sgn}\left\{\pi_2\left(\eta\left((u_1,w_1^0);\Gamma_{g_1^0}\right)\right)\right\}$$
$$=\mathrm{sgn}\left\{g_1^0\left(a\right)-u_1\right\}\cdot or\left(\varphi_1^0\right)$$
$$=\mathrm{sgn}\left\{a_0-u_1\right\}\cdot(+1)$$
$$=(-1)(+1)=-1$$

Therefore P_1^0 is exposed in the direction of $-e_2$, as depicted in Fig. 3 and the ray r_u starting at t_1^0 must intersect $\partial F(R)$ at some point $r_u(t)$ with $t>t_1^0$. If this case does not happen, for notation purposes, let us set $t_1^0=0$.

Suppose there is $t \geq 0$ such that $r_u(t) \in \varphi_1^1$. Let $t_1^1 = \max\{t \mid r_u(t) \in \varphi_1^1\}$ and denote the point $P_1^1 = r_u(t_1^1) = (u_1, w_1^1)$. As we consider φ_1^1 as a graph of a function over the x_2-axis, let us denote this map by $g_1^1 : [a,b] \to \mathbb{R}$, with $g_1^1(a) = a_1$ as depicted in Fig. 4.

First notice that $w_1^1 \in (a,b)$ and from the choice of t_1^0, we have that $w_1^1 = \min\{(g_1^1)^{-1}(u_1)\}$. Observe that the orientation of φ_1^1 is the canonical positive orientation as F normally assembles R. Then, an application of Lemma 2.3 yields the following:

$$\text{sgn}\{\pi_2(\eta(P_1^1; \varphi_1^1))\} = \text{sgn}\{\pi_2(\eta((u_1, w_1^1); \Gamma_{g_1^1}))\}$$
$$= \text{sgn}\{g_1^1(a) - u_1\} \cdot or(\varphi_1^1)$$
$$= \text{sgn}\{a_1 - u_1\} \cdot (+1)$$
$$= (-1)(+1) = -1$$

Therefore P_1^1 is exposed in the direction of $-e_2$, as depicted in Fig. 4 and the ray r_u starting at t_1^1 must intersect $\partial F(R)$ at some point $r_u(t)$ with $t > t_1^1$. If this case does not happen, for notation purposes, let us set $t_1^1 = 0$.

Now, since at least one of the cases above hold, consider $t^* = \max\{t_1^0, t_1^1, t_2^1\} > 0$. Indeed, if one of corresponding case does not hold then we have that $t_j^{\delta_j} = 0$, but by assumption not all three can be zero.

In any case, the point $r_u(t^*)$ is exposed in the direction $-e_2$, thus for some $\tau > t^* \geq 0$ $r_u(\tau) \in \partial F(R)$. Since F is a local diffeomorphism, the image of the boundary of R must contained in the boundary of the image of R, that is, $F(\partial F(R) \subseteq \partial F(R)$, see (Balreira *et al.*, 2014, Lemma 4.4] for a detailed proof. Hence, it must be that $r_u(\tau) \in (\varphi_2^1, \varphi_1^0, \varphi_1^1)$, a contradiction to the choice of t^*.

Therefore, for each $u \in \varphi_2^0$, we have that u cannot belong to any other edge of the boundary, so $F|_{\partial R}$ must be injective and by Theorem 2.1 F is injective.

Finally, we observe that the geometric condition of Theorem 1.1, that is, that F normally assembles the unit square can be easily checked analytically. In fact, we have the following result.

Proposition 2.4

Let R be a rectangle and $F = (f, g) : R \to \mathbb{R}^2$ be a C^1 local diffeomorphism. Suppose that for every $(x, y) \in \partial R$, the Jacobian matrix, $DF(x, y)$ is a P-matrix. Then F normally cubically assembles R.
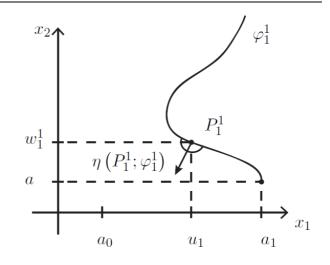

Fig. 4: Intersection between ray $r_u(t)$ and φ_1^1 creating an exposed point P_1^1

We remark that it is sufficient to verify the analytic conditions of $DF(x)$ being a P-matrix on the boundary only, similarly to the results by Garcia and Zangwill (1979).

Proof

It will suffice to show that for $R_2^0 = \{(x, 0) \mid 0 \leq x \leq 1\}$ and $F(R_2^0) = \varphi_2^0$, we have:

$$\langle \eta(P, \varphi_2^0), e_2 \rangle > 0.$$

Indeed, let $\gamma(t) = (f(t, 0), g(t, 0))$, then for $P = (t) \in \varphi_2^0$, we have:

$$\eta(P, \varphi_2^0) = R_{90}(\gamma'(t)) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_x(t, 0) \\ g_x(t, 0) \end{pmatrix}$$
$$= (-g_x(t, 0), f_x(t, 0))$$

Therefore, $\langle \eta(P, \varphi_2^0), e_2 \rangle = f_x(t, 0) > 0$ as DF is a P-matrix.

Conclusion

The geometric interpretation of injectivity provides insights for the proof of the Gale-Nikaidô conjerure in higher dimensions. The introduction of the concept of exposed points have been useful in other problems on injectivity and in problems in discrete dynamics. In article, it has played an important role in determining injectivity and the proof of a generalization of the planar Gale-Nikaidô Theorem.

Acknowledgment

Ethics

The author declare that there is no conflict interests regarding the publication of this manuscript.

References

Aleksandrov, V.A., 1994. On the fundamental Gale-Nikaidô-Inada theorem on the injectivity of mappings. Sibirsk. Mat. Zh., 35: 715-718.

Balreira, E.C., S. Elaydi and R. Luıs, 2014. Local stability implies global stability for the planar Ricker competition model, Discrete Continuous Dynamical Syst., 19: 323-351.
DOI: 10.3934/dcdsb.2014.19.323

Fujimoto, T. and C. Herrero, 2000. A univalence theorem for nonlinear mappings: An elementary approach. Okayama Economic Rev., 31: 277-283.

Gale, D. and H. Nikaidô, 1965. The Jacobian matrix and global univalence of mappings. Math. Ann., 159: 81-93. DOI: 10.1007/BF01360282

Garcıa, C.B. and W.I. Zangwill, 1979. On univalence and P-matrices. Linear Algebra Appl., 24: 239-250. DOI: 10.1016/0024-3795(79)90161-7

Kestelman, H., 1971. Classroom notes: Mappings with non-vanishing jacobian. Am. Math. Monthly, 78: 662-663. DOI: 10.1080/00029890.1971.11992827

Mas-Colell, A., 1979. Two Propositions on the Global Univalence of Systems of Cost Function. In: General Equilibrium, Growth and Trade, Academic Press, New York, pp: 323-331.
DOI: 10.1016/B978-0-12-298750-2.50022-8

Parthasarathy, T., 1983. On global univalence theorems. Springer-Verlag, Berlin-Heidelberg-New York.

Samuelson, P.A., 1949. International factor-price equalisation once again. Economic J., 59: 181-197. DOI: 10.2307/2226683

Samuelson, P.A., 1967. Summary on factor-price equalization. Int. Economic Rev., 8: 286-295. DOI: 10.2307/2525536