Proceedings of the 1998 IEEE/RSJ
Intl. Conference on Intelligent Robots and Systems
Victoria, B.C., Canada • October 1998

Properties of Minimum Infinity-Norm Optimization Applied to Kinematically Redundant Robots

Ian Gravagne and Ian D. Walker
Electrical and Computer Engineering, Clemson University, Clemson SC 29634
igravag@clemson.edu, ianw@ces.clemson.edu

Abstract

There are numerous situations in robotics where it becomes desirable to minimize the maximum magnitude of a solution to an under-determined set of linear equations. For example, there have been several approaches to finding the joint velocities of kinematically redundant robots using this philosophy. Unfortunately, the solution of this optimization problem, known as the "minimum infinity-norm solution", cannot be expressed in a closed form in general, thus requiring the use of an algorithm to iteratively refine an initial guess before reaching the desired solution. In order to increase our understanding and reduce the complexity of infinity-norm algorithms, we first formulate a new "infinity inverse", and then use the new inverse to explore critical issues such as uniqueness and continuity of least infinity-norm solutions. The new inverse is compared with the well-known pseudoinverse, or minimum two-norm solution. We discuss when and why one particular norm might produce a better solution than the other, reinforcing the discussion with an interesting example of a kinematically redundant manipulator.

1 Introduction

One need not ponder too long to find a plethora of situations requiring the optimization of a finite set of variables in terms of some kind of scalar constraint. These applications range from financial investment to airliner flightpath scheduling to manufacturing and production efficiency analysis. Sometimes the tradeoffs can successfully be made in an ad-hoc manner; sometimes there are complex computational tools involved. However, many optimization problems, when phrased mathematically, boil down to a simple problem: the maximization or minimization of a scalar. Plentiful are the techniques for combining all the variables into one measure in a sensible way, but few are the methods for easily optimizing the measure. This reality often rears its head in robotics in the analysis of kinematically redundant robots, when researchers wonder how to best utilize the excess capability of, for example, a 6 degree of freedom (DOF) robot performing a 4 or 5 DOF task, or of a 7 or greater DOF robot in a typical 6 DOF environment. Extra degrees of freedom are highly useful because they allow the execution of subtasks such as obstacle avoidance or singularity avoidance [9,10]. To be specific, redundant kinematic degrees of freedom allow an infinite number of configurations which all accomplish the end-effector trajectory. The end-effector position $\underline{r} \in \Re^m$ relates to the joint angles $\underline{\theta} \in \Re^n$ through the forward kinematic equation

$$\underline{r} = \underline{f}(\underline{\theta}) \tag{1}$$

For redundant robots, given a desired end effector trajectory $\underline{r}(t)$, to find a joint configuration solution $\underline{\theta}(t)$ to generate $\underline{r}(t)$, the velocity level relationship

$$\dot{\underline{r}} = J\dot{\underline{\theta}} \tag{2}$$

is used, where J is the manipulator Jacobian, given by

$$J(\underline{\theta}) = \frac{\partial \underline{f}(\underline{\theta})}{\partial \underline{\theta}}$$

For a kinematically redundant robot, $J \in \Re^{n \times m}$ with $m < n$.

When J is square and non-singular (i.e the non-redundant case), it may simply be inverted to produce $\dot{\underline{\theta}}$ when $\dot{\underline{r}}$ is known. However, in the general case that J is not square (i.e. the robot is redundant), then we must find an optimization measure which allows the determination of a generalized inverse for J. The vast majority of candidate measures ultimately rely on variations of the well-known Moore-Penrose pseudoinverse J^+ [9,10],

$$\dot{\underline{\theta}} = J^+\dot{\underline{r}}; \qquad J^+ = J^T(JJ^T)^{-1} \tag{3}$$

For many applications, the pseudo-inverse is exactly the appropriate answer because it can be modified to minimize the "weighted energy" of the resulting robot motion. However, it is not the only measure of optimization. While pseudo-inverse calculations yield solution vectors of minimum energy, these solutions are only one type in a broad class of optimizations known as minimum norm optimizations. There are, in fact, an infinite number of minimum norm solutions; however, three in particular have easily identified physical meaning. The pseudo-inverse optimizes according to the "2-norm". Cadzow [2,3] explored algorithms for minimizing the "1-norm" and the "infinity-norm". Walker and Deo [7,8] successfully applied the least infinity norm to kinematic redundancy, where it provides solution vectors with the maximum joint velocity minimized. Yoon and Shim [4] applied a geometric interpretation of the least infinity-norm to the dynamics of a redundant robot. As more and more research surfaces which utilizes the least infinity-norm, it is important to examine its structural properties. As we shall see, it is in many ways not as "nice" as the 2-norm, and great care must be exercised when using it in certain applications. However, as we will see in the following section, there are some strong reasons for preferring the least infinity-norm in some situations.

In this paper, we explore the underlying structure of the infinity-norm solutions, including the continuity and uniqueness of solutions derived for continuously evolving systems typified by robot manipulators. For this purpose, we formulate a new "infinity-inverse" solution and relate and compare it with the pseudoinverse. We then use the structure of our solution to investigate uniqueness and continuity properties. Finally, we illustrate the approach using an example.

2 Background

Although there exists a great deal of literature on the subject of generalized inverses of under-determined linear equations, the main mathematical thrust of that work serves mainly to prove that certain classes of inverse exist, demonstrate their qualities, and perhaps hint at the structure of other inverses besides the Moore-Penrose pseudo-inverse. The pseudo-inverse and its many variants remain attractive because those inverses represent a minimization of the Euclidean length, of the solution vector, or a minimization of the "energy" of the corresponding system. More importantly, simple closed-form expressions exist which permit in-depth examination of the characteristics of minimum two-norm solutions. Given the expression[1]

$$A\underline{x} = \underline{b}; \qquad A \in \Re^{mxn}, \underline{x} \in \Re^{n}, \underline{b} \in \Re^{m}, m < n \quad (4)$$

[1]Hereafter, capital letters denote matrices, and underscored small letters denote vectors. If A is a matrix, $\underline{a}_2$ is the 2nd column in A. If $\underline{r}$ is a vector, $\underline{r}(3)$ is the 3rd element of $\underline{r}$.

there are an infinite number of generalized inverses [1] which will produce solutions of the form

$$\underline{x} = A^{*}\underline{b} \quad (5)$$

In the case of the pseudo-inverse, $A^{*} \equiv A^{+} = D^{-1}A^{T}(AD^{-1}A^{T})^{-1}$, and equation (5) solves

$$\min \left\| D^{1/2}\underline{x} \right\|_2$$

subject to

$$A\underline{x} = \underline{b}, \qquad D \in \Re^{nxn}. \quad (6)$$

and D is nonsingular. Sometimes a different performance measure can provide solutions more suited to the physical situation at hand. Recall the definition of a vector p-norm,

$$\left\|\underline{x}\right\|_p = \left(|\underline{x}(1)|^p + |\underline{x}(2)|^p + ... + |\underline{x}(n)|^p\right)^{\frac{1}{p}} \quad (7)$$

so that, in the limit as p goes toward infinity,

$$\left\|\underline{x}\right\|_\infty = \max\left\{|\underline{x}(1)|, |\underline{x}(2)|, ..., |\underline{x}(n)|\right\}. \quad (8)$$

Another useful norm which will appear in conjunction with the infinity-norm is the one-norm, or

$$\left\|\underline{x}\right\|_1 = |\underline{x}(1)| + |\underline{x}(2)| + ... + |\underline{x}(n)|. \quad (9)$$

So, it seems reasonable to re-phrase problem (6) using (8) or (9) as a performance measure to be minimized. The use of (8) as a performance measure yields a solution vector with the minimum maximal magnitude possible, while solving for the two-norm does not guarantee minimum vector component magnitudes, especially in the case that m is much smaller than n. Thus, in cases where individual joint velocities or torques are of concern, the least infinity-norm solution presents an attractive alternative to the two-norm solution [7].

The question naturally arises, how will a least infinity-norm solution behave as a system and its inputs change with time? In other words, will there be any surprises when computing the solution to

$$\min \left\| D^{1/2}\underline{x}(t) \right\|_\infty$$

subject to

$$A(t)\underline{x}(t) = \underline{b}(t). \quad (10)$$

The majority of candidate algorithms for computing the infinity-norm solutions do not permit easy analysis of questions such as sensitivity, continuity and uniqueness in a time-varying situation (issues easily explored in the two-norm solution). These algorithms have been designed for computational efficiency, which usually means many lines of "if-then" conditionals, comparisons, loops and auxiliary computations before a solution arrives [2,3]; possible sources for strange or unexpected behavior might remain difficult if not impossible to pin down. Previous work in infinity-norm solutions in robotics [4,7,8,12] has largely neglected such issues of uniqueness or continuity of solutions. In the following, we develop a generalized inverse for A which solves (10). While not amenable to direct computation, it more importantly yields some insights into the nature of the questions posed by (10).

3 A New Generalized Inverse

It remains unlikely that a generalized inverse will be found which is as simple to compute as the pseudo-inverse. However, a good starting point to begin searching would be the basic geometric arguments which define the norm. Yoon and Shim [4] demonstrated that the least infinity norm has a simple geometric representation, and took advantage of that simplicity to augment the constraint conditions for solving (10). We begin by summarizing the procedure in [4]. Let the set Q represent extrema for an n-dimensional hypercube with unit dimensions, i.e. Q is given by $\{\underline{x} \mid \|\underline{x}\|_\infty = 1\}$. (To prevent carrying D throughout the analysis, let $\underline{x} = D^{1/2}\hat{\underline{x}}$ and $A = \hat{A}D^{-1/2}$ where $\hat{\underline{x}} \in \Re^n$ solves $\hat{A}\hat{\underline{x}} = \underline{b}$.) Set Q really contains an infinite number of elements, but just as an ellipsoid may be characterized by its major axes, a polyhedron's matrix representation needs only the location of each vertex point. So we may declare the vertex elements in Q to be the columns of matrix Q, and two properties of Q make themselves immediately evident: each column will consist of a vector of the form $[\pm 1, \pm 1 \ldots \pm 1]$ representing the corners of the n-dimensional hypercube $\|\underline{x}\|_\infty = 1$, and its dimension is $Q \in \Re^{n \times 2^n}$. Then, A maps the points in Q to a m-dimensional, closed, convex polyhedron Γ whose vertex points are contained in the set P. Again, letting each element of P serve as a column of the matrix P, $P \in \Re^{m \times 2^n}$. To summarize this mapping,

$$\begin{aligned} A : Q &\longrightarrow P \\ \Longrightarrow AQ &= P \end{aligned} \tag{11}$$

In most cases, the mapping of vertices is bijective (i.e. onto and one-to-one), but we will see that certain conditions can cause the mapping to lose bijectivity. Now a theorem from convex set theory [5] proves useful:

Theorem 1. *Given the matrix V where the columns $\underline{v}_i$ $(i = 1, 2, \ldots, K)$ represent the m-tuple vertices of a closed, bounded convex polyhedron Γ, all extreme boundary points $\underline{p}_B$ of Γ must be represented as a convex combination of those vertices,*

$$\underline{p}_B = V\underline{\mu}; \qquad \underline{\mu} \in \Re^K, \mu_i \geq 0,$$

and

$$\sum_{i=1}^{K} \mu_i = 1. \tag{12}$$

If $\underline{v}_B$ is not a boundary point, but rather an interior point in Γ, then there always exists some $\underline{\mu}$ such that $\sum_{i=1}^{K} \mu_i < 1$.

Conversely, we may formalize the definition of Γ by noting that a closed polyhedron contains the set of all points $\underline{p}$ which can be represented as combinations of the vertex points which are columns in V.

$$\Gamma = \{\underline{p} \mid \underline{p} = V\underline{\mu} : \sum_{i=1}^{K} \mu_i \leq 1\} \tag{13}$$

The surface of Γ is covered by hyperplanes containing only those $\underline{p}$ where equality holds in (13). To deduce the least infinity norm solution to (4), first find the hyperplane G on the boundary of Γ such that the vector $\lambda\underline{b}$ intersects G with $\lambda \in \Re$. In other words, scale vector $\underline{b}$ until its endpoint belongs to plane G. Note it is possible for $\lambda\underline{b}$ to point toward the surface where several hyperplanes converge, but in general the intersection of k m-dimensional hyperplanes is a (m-k+1)-dimensional hyperplane. This is the analog of two planes converging to a line, or three planes converging to a point. To simplify the discussion, let us assume there is only one intersecting hyperplane G. G is defined by exactly m vertices on Γ, each with a corresponding column in P. Thus, the sub-matrix containing only those vertices is labeled $P_{i \in G}$ and the points $\underline{q}$ which generated those vertices form the sub-matrix $Q_{i \in G}$, so $AQ_{i \in G} = P_{i \in G}$. When the tip of $\lambda\underline{b}$ meets hyperplane G, $\lambda\underline{b}$ is a boundary point and therefore a convex combination of the columns of $P_{i \in G}$, so there is at least one $\underline{\mu}$ which satisfies the conditions in Theorem 1, i.e. $\lambda\underline{b} = P_{i \in G}\underline{\mu}$. Remembering that $\underline{b} = A\underline{x}$ and $P_{i \in G} = AQ_{i \in G}$, we may say $\lambda A\underline{x} = AQ_{i \in G}\underline{\mu}$. Thus, when $\lambda, \underline{\mu}$ and $Q_{i \in G}$ are known, there is a particular solution where

$$\underline{x}^* = \frac{1}{\lambda} Q_{i \in G}\underline{\mu}. \tag{14}$$

$\underline{x}^*$ is in fact the least-infinity norm solution to (4) [4]. Unfortunately, finding the particular vertices which form the hyperplane intersecting $\lambda\underline{b}$ can prove exceedingly time-consuming because not every element of P is a boundary point on Γ; especially if $m \ll n$, a large number of those elements are interior points.

At this point, most previous discussions of the least-infinity norm structure either proceed to direct computation λ and $\underline{\mu}$, or begin to identify strategies for iteratively determining the solution. However, there has been little discussion or investigation of the structure of the iterative solution. In the following we explore these issues more closely. In this paper, we are not so concerned with computation as with structure, so we first multiply P by a diagonal weighting matrix $W \in \Re^{2^n \times 2^n}$ where the numbers on the diagonal represent relative "probabilities" that the corresponding column of P belongs to hyperplane G. According to Theorem 1, the point at which $\lambda \underline{b}$ intersects G must be a convex combination of vertices $P_{i \in G}$ so incorporating the "guessing" matrix W gives

$$PW\underline{\tilde{\mu}} = \tilde{\lambda}\underline{b} \tag{15}$$

Here $\underline{\tilde{\mu}}$ and $\tilde{\lambda}$ are simply the estimated versions of $\underline{\mu}$ and λ. From this point, Theorem 2 follows.

Theorem 2. *Given the underdetermined linear expression (4), with Q and P as defined above, the least infinity norm solution $\underline{x}^*$ can be written as*

$$\underline{x}^* = Q_{i \in G} \left(AQ_{i \in G}\right)^{-1} \underline{b} \tag{16}$$

Proof Starting from (15), we would like to know at least one possible $\underline{\tilde{\mu}}$ which satisfies the expression, so let

$$\underline{\tilde{\mu}} = (PW)^{+}\tilde{\lambda}\underline{b} \tag{17}$$

In general the pseudo-inverse will not yield a nonnegative $\underline{\tilde{\mu}}$ whose elements sum to 1, (i.e. satisfying Theorem 1), but this can be corrected by assuming W is known well enough to produce $\underline{\tilde{\mu}}$ with $\tilde{\mu}_i \geq 0$, and then normalizing $\underline{\mu} = \frac{\underline{\tilde{\mu}}}{\|\underline{\tilde{\mu}}\|_1}$, and $\lambda = \frac{\tilde{\lambda}}{\|\underline{\tilde{\mu}}\|_1}$ so that $\sum_{i=1}^{2^n} \mu_i = 1$. Now

$$PW\underline{\mu} = \lambda\underline{b} \Longrightarrow \underline{\mu} = (PW)^{+}\lambda\underline{b} \tag{18}$$

With $\underline{\mu}$ now conforming to Theorem 1, the corresponding inverse mapping gives

$$QW\underline{\mu} = \lambda\underline{x} \tag{19}$$

Rearranging and substituting for $\underline{\mu}$ and λ, (19) becomes $\underline{x} = \frac{1}{\lambda}QW\underline{\tilde{\mu}}$. Substituting for $\underline{\tilde{\mu}}$ from (17),

$$\underline{x} = QW(PW)^{+}\underline{b} \tag{20}$$

If W is known exactly, we may eliminate unused columns of P, and (20) reduces to the desired result

$$\underline{x}^* = Q_{i \in G}P_{i \in G}^{-1}\underline{b} = Q_{i \in G}(AQ_{i \in G})^{-1}\underline{b} \tag{21}$$

Q.E.D.

The above result provides structure to the infinity-norm solution, which allows us to relate it to the much used pseudoinverse, or 2-norm solution as follows.

Relaxing the initial assumptions about W yields an interesting observation: if W is not known at all (all entries on the diagonal are equal and constant), then (20) may be rewritten[2] as

$$\begin{aligned} \underline{x}^* = QWW^{+}P^{+}\underline{b} &= QWW^{-1}(AQ)^{+}\underline{b} \\ &= QQ^{+}A^{+}\underline{b} = A^{+}\underline{b} \end{aligned} \tag{22}$$

In this manner, the least two-norm and least infinity-norm solutions directly relate to each other based only upon how well W is known. Henceforth, the coefficient to $\underline{b}$ in (20) introduced here will be called the "infinity-inverse", labelled $A^{\#}$. $A^{\#}$ is a generalized inverse, of a class sometimes referenced as a {1,2,3} inverse. The numbers in brackets denote which of the four Moore-Penrose conditions a generalized inverse obeys; the pseudo-inverse is the only {1,2,3,4} inverse. Note that the structure of $A^{\#}$ closely matches the structure of A^{+}; the dimension of $Q_{i \in G}$ is the dimension of A^T. The Moore-Penrose Conditions are given below:

$$\begin{array}{ll} 1) AXA = A & 2) XAX = X \\ 3) (AX)^T = AX & 4) (XA)^T = XA \end{array}$$

X above signifies any potential generalized inverse. For $X = A^{\#}$, conditions 1 and 2 are trivially proven (except under rare conditions to be illustrated in the next section); condition 3 follows because $(AA^{\#})^T = (AQ_{i \in G}(AQ_{i \in G})^{-1})^T = I$. Simultaneously satisfying conditions 1 and 2, and at least one of either 3 or 4, allows a generalized inverse to operate as an orthogonal projector on the nullspace of A [1,6]. Thus, the most general form of solution to (1) holds for $A^{\#}$ just as for A^{+}:

$$\underline{x} = A^{\#}\underline{b} + (I - A^{\#}A)\underline{\varepsilon} \tag{23}$$

and $\underline{\varepsilon} \in \Re^n$ is arbitrary. Given the apparent symmetry between (23) and the same equation using the pseudoinverse, it can be observed that the 2-norm solution can be derived by an infinity-inverse summed with an appropriate nullspace offset. Similarly, the least infinity-norm solution may be derived with a pseudoinverse added to the negative of the same nullspace offset. However, the infinity-inverse is not merely a weighted pseudoinverse, and vice-versa.

The symmetry does not end there. Heretofore and hereafter, we assume full rank systems, however the

[2] It is difficult to predict he exact conditions under which the pseudo-inverse of a matrix product equates to the product of the pseudo-inverses. One thing is known, however [1]: if $r(A^TAQ) \subset r(Q)$ and $r(QQ^TA^T) \subset r(A^T)$, then $(AQ)^+ = Q^+A^+$, where $r(A)$ denotes the rowspace of A. Given the nature of the columns of Q, this result holds for all compatible A, though the proof is not shown here.

infinity-inverse will behave similarly to the pseudoinverse in the overdetermined case: where the pseudoinverse yields the 'least mean squared error' solution, the infinity-inverse will yield the 'least error magnitude' solution.

4 Conditions for Injective Failure

We should recognize a few important observations regarding the geometry at hand. As mentioned earlier, usually the mapping A is bijective. Points which define a particular boundary hyperplane G on Γ in m-space are "nearest neighbors" on the hypercube in n-space [11]. Given a point in n-space $\underline{q}_0$ the set of nearest neighbors contains n vectors $\underline{q}_{i=1,2,..,n}$ each with scalar components

$$\underline{q}_i(j) = \left\{ \begin{array}{ll} -\underline{q}_0(j) & j = i \\ \underline{q}_0(j) & j \neq i \end{array} \right\} \qquad j = 1, 2, ..., n \tag{24}$$

Recall that Q consists only of 1 and -1, so for example, if $\underline{q}_0 = [\ 1 \quad -1 \quad 1\ ]^T$, then the complete set of nearest neighbors would be $\underline{q}_1 = [\ -1 \quad -1 \quad 1\ ]^T$, $\underline{q}_2 = [\ 1 \quad 1 \quad 1\ ]^T$, and $\underline{q}_3 = [\ 1 \quad -1 \quad -1\ ]^T$. Assuming $A \in \Re^{2\times 3}$, and observing that $[\ \underline{p}_0 \quad \underline{p}_1 \quad \underline{p}_2 \quad \underline{p}_3\ ] = A[\ \underline{q}_0 \quad \underline{q}_2 \quad \underline{q}_3 \quad \underline{q}_4\ ]$, then $\underline{p}_1$, $\underline{p}_2$ and $\underline{p}_3$ are nearest neighbors to $\underline{p}_0$. Matrix A will probably not map every $\underline{p}_i$ to a vertex in m-space; some will map to interior points of the polyhedron. However, if a given $\underline{p}_0$ does locate on a vertex, each associated boundary hyperplane of which $\underline{p}_0$ is a member will consist of at least m-1 nearest neighbors to $\underline{p}_0$ because A is a linear transformation. It is important to remember that the measure of "nearest" in m-space is not based on a Euclidian metric, but rather the notion that nearest-neighbor points in m-space directly descend from Euclidian nearest-neighbors in n-space. So, every face G of Γ is composed of vertices descended from nearest neighbors in Q, meaning that every column of $Q_{i\in G}$ must be a nearest neighbor to every other column and therefore independent of every other column. With this information in hand, we may address the question of when transformation A ceases to be injective, so that distinct elements of Q map to identical elements of P. One more semantic clarification will help: when saying, "k columns of matrix A sum to equal another column of A", the meaning is actually "one column of A is a linear combination of k other columns, with coefficients 1 or -1 only." In other words, some columns may be subtracted while others are added; we generically refer to this arithmetic combination as a column sum.

Theorem 3 *Define $A \in \Re^{m\times n}$ with $m < n$ and $rank(A) = m$, as in (4). If two or more columns of A sum to equal any other column of A then the mapping $A : Q \longrightarrow P$ is not one-to-one given Q as $\{\underline{q} \mid \|\underline{q}\|_\infty = 1 \text{ with } \underline{q} \in \Re^n\}$.*

Proof Partition A so $A = [B, C, \underline{a}_k]$. $C \in \Re^{m\times c}$, and $B \in \Re^{m\times(n-c-1)}$. Let the columns of C sum to equal column $\underline{a}_k$, so

$$C\underline{\alpha} = \underline{a}_k \qquad \underline{\alpha}(j) \in \{-1, 1\} for j = 1, 2, ..., c \tag{25}$$

This is equivalent to the conditions given in the theorem statement. Define a column of Q, say, $\underline{q}_0 = \left[\underline{\beta}^T, \underline{\alpha}^T, -1\right]^T$ with $\underline{\beta}(i) \in \{-1, 1\}$ for $i = 1, 2, ...(n - c - 1)$. Pre-multiplying by A so that $\underline{p}_0 = A\underline{q}_0$ and simplifying by (25) gives

$$\begin{aligned} \underline{p}_0 &= [\ B \quad C \quad \underline{a}_k\ ] \begin{bmatrix} \underline{\beta} \\ \underline{\alpha} \\ -1 \end{bmatrix} \\ &= \left[B\underline{\beta} + C\underline{\alpha} - \underline{a}_k\right] = \left[B\underline{\beta}\right]. \end{aligned} \tag{26}$$

However, defining $\underline{q}_1 = \left[\underline{\beta}^T, -\underline{\alpha}^T, 1\right]^T$, pre-multiplying by A and simplifying also yields

$$\begin{aligned} \underline{p}_1 &= [\ B \quad C \quad \underline{a}_k\ ] \begin{bmatrix} \underline{\beta} \\ -\underline{\alpha} \\ 1 \end{bmatrix} \\ &= \left[B\underline{\beta} - C\underline{\alpha} + \underline{a}_k\right] = \left[B\underline{\beta}\right] = \underline{p}_0. \end{aligned} \tag{27}$$

So, the result follows that if (25) holds for any columns A, at least two columns of P will be identical, so mapping A ceases to be injective (one-to-one), and therefore not bijective. Q.E.D.

Remark Theorem 3 will hold not only for columns of A which sum to another column A, but also whenever a column of A equals zero. Except in the case of a null column, equally important is the fact that the offending columns of Q, (i.e. $\underline{q}_0$ and $\underline{q}_1$) satisfying (26) and (27), cannot be nearest neighbors in n-space despite that fact that $\underline{p}_0$ and $\underline{p}_1$ are identical in m-space.

Actually, the case that a column of A equals zero is the only condition which would allow $\underline{q}_0$ and $\underline{q}_1$ to be nearest neighbors. In this rare instance, a m-dimensional hyperplane face G (on polyhedron Γ) which was defined by m independent vertices $\underline{p}$ will now shrink to an (m-1)-dimensional face still defined by m vertices, two of which are identical and all of which are nearest neighbors. In effect, $rank(AQ_{i\in G}) < m$ so $(AQ_{i\in G})^{-1}$ will not exist. For $AQ_{i\in G}$ to lose rank, and for $\underline{b}$ to point toward G exactly at that moment presents an extremely unlikely situation, but it is nevertheless a situation where the infinity-inverse will fail even if A has full rank.

As we will see below, problems with uniqueness and continuity of solutions, which are critical issues for the practicality of using the infinity-norm approach, are strongly related to injective failure.

5 Uniqueness

The question of uniqueness of least infinity norm solutions should be discussed in two distinct cases. The first is the simple case that vector $\lambda \underline{b}$ intersects only one hyperplane G on the boundary of Γ in m-space. If A is full rank and provides a one-to-one mapping, $AQ_{i\in G}$ is non-singular and unique so $(AQ_{i\in G})^{-1}$ exists and is unique, thus making $A^{\#}$ unique. The second case requires investigation into what occurs at hyper-plane intersections and vertices in m-space. Two intersecting faces of a convex polyhedron, each defined by at least m vertices, must share at most m-1 vertices. To formalize, let G be the set of all points in a hyperplane defined by vertices $P = \left[\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_{i-1}, \underline{p}_i, \underline{p}_{i+1}, \ldots, \underline{p}_m\right]$. Let $\widetilde{G}$ be the set of all points in an intersecting hyperplane defined by vertices $\widetilde{P} = \left[\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_{i-1}, \widetilde{\underline{p}}_i, \underline{p}_{i+1}, \ldots, \underline{p}_m\right]$, where $\widetilde{\underline{p}}_i \neq \underline{p}_i$, $AQ_{i\in G} = P$ and $A\widetilde{Q}_{i\in\widetilde{G}} = \widetilde{P}$. Define $\underline{b}$ as some linear combination of points in the intersection plane $G \cap \widetilde{G}$ (of dimension m-1), i.e.

$$\underline{b} = \widetilde{P}\underline{\gamma} = P\underline{\gamma} \qquad \underline{\gamma}(i) = 0$$

and

$$\underline{\gamma} \in \Re^m. \tag{28}$$

Since $\lambda \underline{b} \in (G \cap \widetilde{G})$, then $\lambda \underline{b} \in G$ and $\lambda \underline{b} \in \widetilde{G}$, therefore $\underline{x} = QP^{-1}\underline{b}$ and $\widetilde{\underline{x}} = \widetilde{Q}\widetilde{P}^{-1}\underline{b}$ are both valid, least infinity norm solutions of $A\underline{x} = \underline{b}$ and $A\widetilde{\underline{x}} = \underline{b}$, respectively. From (22),

$$\widetilde{\underline{x}} = \widetilde{Q}\widetilde{P}^{-1}\underline{b} = \widetilde{Q}\widetilde{P}^{-1}\widetilde{P}\underline{\gamma} = \widetilde{Q}\underline{\gamma}$$

and

$$\underline{x} = QP^{-1}\underline{b} = QP^{-1}P\underline{\gamma} = Q\underline{\gamma} \tag{29}$$

but since $\underline{\gamma}(i) = 0$, column $\widetilde{\underline{q}}_i$ does not affect $\widetilde{\underline{x}}$, and $\underline{q}_i$ does not affect $\underline{x}$. $Q_{i\in G}$ differs from $\widetilde{Q}_{i\in\widetilde{G}}$ only in the i^{th} column, so $\widetilde{\underline{x}} = \underline{x}$, implying that least infinity norm solutions remain unique even at the intersection of two (or more, though not shown here) faces of Γ.

The problem arises, naturally, when A behaves as in Theorem 3, and the mapping $AQ = P$ is not one-to-one. Recall from an earlier discussion that the vertices defining a face of polyhedron Γ must be nearest neighbors. Usually, this implies that the intersection of two faces would also be comprised of nearest neighbors. However, when A is not injective, while each individual face of Γ still consists of nearest neighbor vertices, the intersection of two faces may not consist of nearest neighbors because a single vertex point in the intersection of two faces could descend from two distinct vertices on the hyper-cube in n-space. Again, define $P = \left[\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_{i-1}, \underline{p}_i, \underline{p}_{i+1}, \ldots, \underline{p}_{j-1}, \underline{p}_j, \underline{p}_{j+1}, \ldots, \underline{p}_m\right]$, $\widetilde{P} = \left[\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_{i-1}, \widetilde{\underline{p}}_i, \underline{p}_{i+1}, \ldots, \underline{p}_{j-1}, \widetilde{\underline{p}}_j, \underline{p}_{j+1}, \ldots, \underline{p}_m\right]$ with $\widetilde{\underline{p}}_i \neq \underline{p}_i$, but $\widetilde{\underline{p}}_j = \underline{p}_j$. Now assume A is no longer one-to-one so that $\widetilde{\underline{q}}_j \neq \underline{q}_j$ but $A\widetilde{\underline{q}}_j = A\underline{q}_j = \widetilde{\underline{p}}_j = \underline{p}_j$. P differs from $\widetilde{P}$ only in the i^{th} column, but the corresponding $Q_{i\in G}$ and $\widetilde{Q}_{i\in\widetilde{G}}$ now differ in more than one column, $\widetilde{\underline{q}}_i \neq \underline{q}_i$, and $\widetilde{\underline{q}}_j \neq \underline{q}_j$. Defining $\underline{b}$ and $\underline{\gamma}$ as in (27), note $\underline{\gamma}(j) \neq 0$, so

$$\widetilde{\underline{x}} = \widetilde{Q}\underline{\gamma} \neq \underline{x} = Q\underline{\gamma} \tag{30}$$

even though $\widetilde{\underline{x}}$ and $\underline{x}$ are both valid, least infinity norm solutions to $A\underline{x} = \underline{b}$ and $A\widetilde{\underline{x}} = \underline{b}$. In fact, in this situation where A fits the characteristics of Theorem 3, $\underline{b}$ is a linear combination of points in $G \cap \widetilde{G}$, and $G \cap \widetilde{G}$ contains at least one non-unique vertex point, then the elements of $\underline{x}^*$ corresponding to the columns of C (and column $\underline{a}_k$) may be assigned any magnitude less or equal to than $\|\underline{x}^*\|_\infty$, including zero. All vectors $\underline{x}^*$ fitting this description are perfectly valid least infinity norm solutions to (4). In other words, with C as in Theorem 3, under these conditions there are an infinite number of least-infinity-norm solutions, with the only requirement $-\|\underline{x}^*\|_\infty \leq \underline{x}^*(i \in C, i = k) \leq \|\underline{x}^*\|_\infty$ (see Theorem 3). The minimum infinity-norm solution to (4) is thus not unique under all conditions.

6 Continuity

With knowledge of how and why a particular solution might not be unique, a discussion of the continuity of such solutions in a continuous-time environment follows easily. Recall that $\underline{x}^* = Q_{i\in G}P^{-1}_{i\in G}\underline{b} = Q_{i\in G}(AQ_{i\in G})^{-1}\underline{b}$. (For the sake of brevity, we will drop the '${}_{i\in G}$' subscript in this section). For this function to exhibit continuity, there must always exist some small $\delta\underline{b}$ and δA which can effect an arbitrarily small change in $\underline{x}$. Using the chain rule several times,

$$\begin{aligned}\delta\underline{x} &= QP^{-1}(\delta\underline{b}) + (\delta Q)P^{-1}\underline{b} \\ &+ Q[(P + (\delta A)Q + A(\delta Q))^{-1} - P^{-1}]\underline{b}\end{aligned} \tag{31}$$

In fact, while $\lambda\underline{b}$ traces out a trajectory along some particular face G of polyhedron Γ, as long as $\underline{b}$ does not cross over to a neighboring face $\widetilde{G}$, then $\delta Q = 0$. In this case, letting $P = AQ$ gives

$$\delta\underline{x} = Q[AQ]^{-1}(\delta\underline{b}) + Q[(AQ + (\delta A)Q)^{-1} - (AQ)^{-1}]\underline{b} \tag{32}$$

and

$$\lim_{\delta\underline{b}\to\underline{0},\delta A\to 0} \delta\underline{x} = \underline{0} \tag{33}$$

so, adequately small changes in $\underline{b}$ and A can produce arbitrarily small changes in $\underline{x}$. Provided A and $\underline{b}$ both evolve continuously, $\underline{x}$ will too. If $\lambda\underline{b}$ crosses a boundary between two or more faces on Γ, corresponding columns of P, and therefore of Q, will abruptly change, so there will be non-zero components in those columns of δQ. However, as discussed in the section on uniqueness, if A is bijective, the corresponding elements of the vector $[AQ]^{-1}\underline{b}$ and $[A(\delta Q)]^{-1}\underline{b}$ will equal zero during this change, so expression (33) holds for all continuous $\underline{b}$ with continuous, bijective A. On the other hand, if $\lambda\underline{b}$ crosses boundary $G\cap\widetilde{G}$ and the boundary plane contains a non-unique vertex point, then $\delta Q \neq 0$, and expression (31) shows that

$$\lim_{\delta\underline{b}\to\underline{0},\delta A\to 0} \delta\underline{x} = (\delta Q)[AQ]^{-1}\underline{b}$$

$$+ Q[(AQ + A(\delta Q))^{-1} - (AQ)^{-1}]\underline{b} \neq \underline{0} \tag{34}$$

implying that the least infinity-norm solution can exhibit discontinuities at boundary-plane intersections with non-bijective A.

7 Discussion and Examples

It might seem as though the conditions for misbehavior of least infinity norm solutions are so rare that only contrived examples would serve to illustrate the associated problems. Nevertheless, the real impetus to search for answers to the question of continuity came when a simple simulation of a 4-DOF planar redundant manipulator began to exhibit some very strange symptoms. The simulator utilized least-infinity norm solutions in the velocity domain. Let $\underline{r}$ denote the end effector position, and $\underline{\theta}$ the angular position of the joints. The velocity of the end effector relates to the angular velocities of each joint as

$$J\dot{\underline{\theta}} = \dot{\underline{r}} \qquad\qquad J \in \Re^{2x4} \tag{35}$$

so, according to (22)

$$\dot{\underline{\theta}} = J^{\#}\dot{\underline{r}} + (I - J^{\#}J)\underline{\varepsilon}.$$

Here, for simplicity, we will not use the manipulator self-motion, letting $\underline{\varepsilon} = \underline{0}$. $J^{\#} = Q_{i\in G}(JQ_{i\in G})^{-1}$ does not lend itself easily to direct computation, because of the need to determine the columns of $Q_{i\in G}$, but the effects are identical when using an algorithm to iteratively find the least infinity-norm solution for $\dot{\underline{\theta}}$.

Of practical concern is the fact that the preceding work applies to the space $\Re^m$, although at most three of a manipulator's variables belong to only a portion of the real line, $[0, 2\pi)$. However, the least norm solutions in the redundant robot context apply only in velocity space, and angular velocities belong to the entire space $\Re^m$.

Before continuing, an important property of minimum infinity-norm solutions should be pointed out: 'Holder's inequality' is a generalization of the Cauchy-Schwarz inequality, and states that

$$\underline{x}^T\underline{b} \leq ||\underline{x}||_p \cdot ||\underline{b}||_q$$

with

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{36}$$

If we let $p \to \infty$, then $q = 1$ and the condition for equality in (31) is known as alignment. If $\underline{b} \in \ell_1$ is given, then

$$\underline{x}^T\underline{b} = ||\underline{x}||_\infty \cdot ||\underline{b}||_1 \iff$$
$$\underline{x}(i) = \left\{ \begin{array}{ll} \alpha \cdot sign[\underline{b}(i)] & \underline{b}(i) \neq 0 \\ \alpha_i & \underline{b}(i) = 0,\ |\alpha_i| < |\alpha| \end{array} \right\}. \tag{37}$$

The famous "duality theorem" [2,7] used in many of the algorithms for finding minimum infinity-norm solutions provides exactly the right framework for using the alignment condition, with the side-effect that $\underline{b}$ will have no fewer than $n - (m-1)$ nonzero components. Stated another way, the optimum infinity-norm solution will solve uniquely for only $m-1$ components of the solution vector; the other $n - m + 1$ components will be identical in magnitude, with that magnitude equal to $\min ||\underline{x}||_\infty$.

Back to the example, the questions arises, what will happen when the end effector traces out a trajectory where the 'best' solution clearly requires that $n - m$ components to go to zero? In the land of 4-DOF planar manipulators, commanding the end-effector to track a circle around the origin requires exactly the aforementioned conditions (figure 1).

The joint required to move at maximum speed is joint 1, the origin. Naturally, it follows that the least infinity-norm will, at each time step, minimize $\dot{\underline{\theta}}(1)$. It works almost too well, though, because in order to compensate for a slower first joint, the manipulator must extend all the subsequent links further and further to achieve the tracking goal. This wouldn't present a problem except for the fact that now only one joint is contributing to the motion of the end-effector. If a pseudo-inverse solution were driving the trajectory, the other 3 joints would simply stop, with magnitude zero. In the infinity-inverse, $n - m + 1$ joints must move at the speed of the fastest joint, joint 1 in this case; they can't simply stop. So why don't the other joints appear to move? The answer becomes clear with a plot of the individual joint velocities (figure 2).

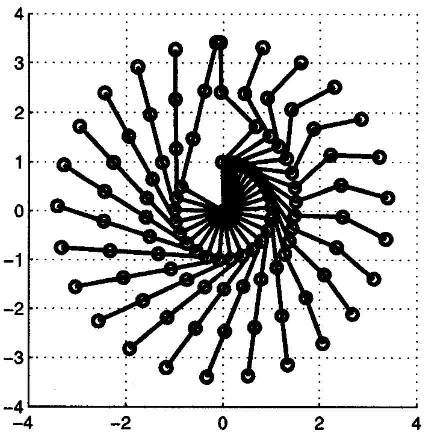

Figure 1: Here the manipulator attempts to trace a circle about the origin. The initial configuration points vertically, with initial angles $\underline{\theta}_0$ [90,45,135,90] in CCW degrees from horizontal. The end effector travels clockwise. Note how the links converge to an 'L'-shaped configuration toward the end of the run.

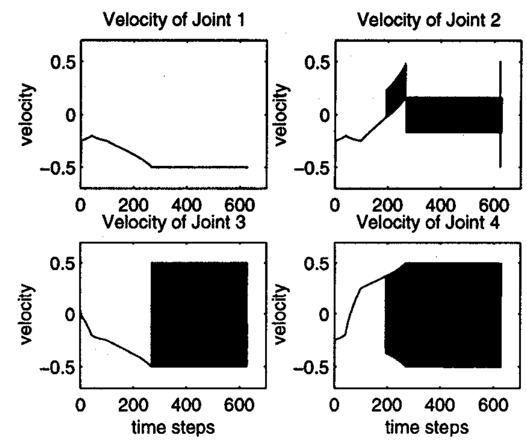

Figure 2: Joint velocities for the circle trajectory.

Approximately the first 200 time steps proceed smoothly while 3 of the 4 joints move at equal speeds. After that point, in order to approximate "zero" motion but still obey the rules of least-infinity norms, joints 3 and 4 rapidly oscillate between the only two values allowed: $-\left\|\dot{\underline{\theta}}\right\|_\infty$ and $\left\|\dot{\underline{\theta}}\right\|_\infty$. Joints 3 and 4 jerk back and forth as rapidly as possible to nullify the average motion over the rest of the trajectory. A total of $(m-1)$ joints are allowed a unique velocity less than $\left\|\dot{\underline{\theta}}\right\|_\infty$ in general, a distinction awarded to joint 2 in this case, although it must attempt to compensate for positional errors in joints 3 and 4 so it also oscillates. From a geometric standpoint, figure 3 illustrates that the Jacobian in this example does indeed meet the criteria of Theorem (3) toward the end of the trajectory, and the desired end-effector velocity vector points very near to a pair of identical vertices.

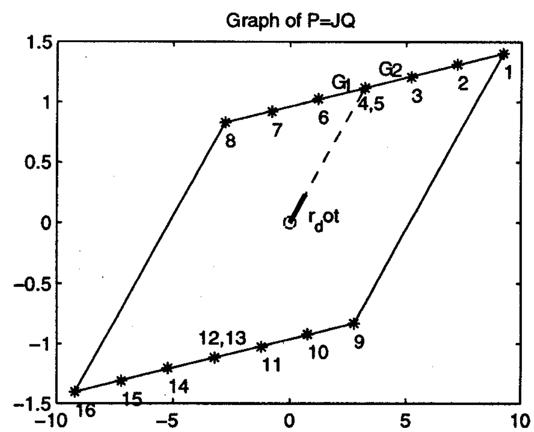

Figure 3: A picture of $P = JQ$. The desired velocity vector is plotted from the origin, labeled r_dot. A dashed line extends its path to intersect exactly with a discontinuity.

An example Jacobian from this trajectory is

$$J = \begin{bmatrix} -3.2263 & -2.9966 & -1.9975 & -0.9990 \\ -1.1166 & -0.1433 & -0.0987 & -0.0454 \end{bmatrix}.$$

Columns three and four add to produce column two. Figure 3 illustrates the convex set $P = JQ$, with $\dot{\underline{r}}$ pointing directly at non-unique vertices $\underline{p}_4$ and $\underline{p}_5$. Polyhedron face G_2 is denoted by vertices $\underline{p}_3$ and $\underline{p}_4$, while face G_1 is composed of vertices $\underline{p}_5$ and $\underline{p}_6$. Note that G_1 cannot be constructed using $\underline{p}_4$ and $\underline{p}_6$ because they are not nearest neighbors, i.e. $\underline{q}_4$ and $\underline{q}_6$ differ by more than one sign.

$$Q = \begin{bmatrix} -1 & -1 & -1 & -1 & & 1 & 1 \\ -1 & -1 & -1 & -1 & & 1 & 1 \\ -1 & -1 & 1 & 1 & \cdots & 1 & 1 \\ -1 & 1 & -1 & 1 & & -1 & 1 \end{bmatrix}$$

Velocity vector $\underline{\dot{r}}$ approaches from face G_1 toward the discontinuity and at iteration k (somewhere around time-step 200 in the velocity graphs) crosses it, pointing to G_2. At k+1, a new Jacobian is computed such that the old $\underline{\dot{r}}$ points again at G_1. At k+2, $\underline{\dot{r}}$ again crosses over to G_2 and a new Jacobian is computed. This cat-and-mouse game will never resolve for a least infinity-norm algorithm. If the algorithm could run in continuous time, at some time $\underline{\dot{r}}$ would point directly at the discontinuity, and the controller could command joints 2, 3 and 4 to zero velocity which would still be a valid least infinity-norm solution. This action would prevent the subsequent oscillations, though the velocities in three joints would still suffer one discontinuity.

8 Conclusions

Minimum infinity-norm solutions are attractive in kinematic redundancy resolution for manipulators due to their potential for explicitly addressing individual joint velocities in the solution. This is in contrast to the 2-norm solution which results from the well-used pseudoinverse technique, where only the effect of a combination of all the joint velocities can be analyzed. However, the structure of the minimum infinity-norm solution is much less well understood than the pseudoinverse solution.

This paper has attempted to expose some of the details and features of minimum infinity-norm solutions to under-determined linear systems, with application to redundancy resolution for kinematically redundant manipulators. The creation of an "infinity-inverse" of a very similar structure to the pseudo-inverse aided immensely in revealing some difficulties with least infinity-norm solutions. The infinity-inverse, however, really does not lend itself as a computational tool because $Q_{i\in G}$ remains difficult to compute, and its relation to $\underline{A}$ and $\underline{b}$ in equation (1) is uncertain. Unfortunately, some of these problems stem from the very nature of the infinity norm. It looks only at the individual components of each solution vector, whereas the two-norm considers the contribution of all components. This characteristic causes non-uniqueness and discontinuity on least infinity-norm solutions.

Prospective users of least infinity-norm algorithms should not lose hope, however, because such solutions have been proven to avoid some of the problems associated with two-norms. Perhaps an intelligent control scheme could switch between infinity-norm and two-norm solutions depending on the conditions. Perhaps least-norm solutions have inherent limitations and some other criteria might serve better.

9 Acknowledgments

The authors are grateful for support from: National Science Foundation grant CMS-9796328; Dept. of Energy grant DE-FG07-97ER14830; Clemson University and the Clemson Electrical and Computer Engineering Department.

References

[1] A. Ben-Israel and T. Greville, Generalized Inverses: Theory and Applications, Pub. Wiley & Sons, 1974

[2] J.A. Cadzow, "Algorithm for the minimum-effort problem," IEEE Trans. Automat. Control., vol AC-16, pp. 60-63, Feb. 1971

[3] J.A. Cadzow, "A finite algorithm for the minimum l_∞-norm solution to a system of consistent linear equations," SIAM Journal of Numerical Analysis, vol. 10, no. 4, pp. 607-617

[4] I.C. Shim and Y.S. Yoon, "Stability constraint for torque optimization of a redundant manipulator," Proc. IEEE Conf. Robotics and Automation, Albuquerque NM, April 1997, pp. 2403-2408

[5] L. Cooper and D. Steinberg, Methods of Optimization, Pub. W.B. Saunders Co., 1970

[6] G. Golub and C.F.Van Loan, Matrix Computations, 3rd ed., pub. Johns Hopkins U. Press, 1996

[7] A.S. Deo and I.D. Walker, "Minimum effort inverse kinematics for redundant manipulators," IEEE Trans. Robotics and Automation, Vol. 13, No. 5, pp. 767-775

[8] A.S. Deo and I.D. Walker, "Methods of redundancy resolution by infinity-norm minimization," Proc. ISRAM '94: Res., Educ., Applicat, ASME Press, NY, 1994, pp. 67-74

[9] D.N. Nenchev, "Redundancy resolution through local optimization: a review," Journal of Robotic Systems, Vol. 6, No. 6, pp. 769-798, 1989

[10] B. Siciliano, "Kinematic control of redundant robot manipulators: a tutorial," Journal of Intelligent and Robotic Systems, No. 3, pp. 201-212, 1990

[11] I. Gravagne, "Properties of minimum infinity-norm optimization applied to continuous systems," Internal Report, Department of Electrical and Computer Engineering, December 1997

[12] B.J. Martin and J.E. Bobrow, "Determination of minimum-effort motions for general open chains," Proc. IEEE Conf. Robotics and Automation, Nagoya, Japan, May 1995, pp. 1160-1165