that is, the matrix $(\boldsymbol{R}^*_{m,j})^+$ can be obtained from the matrix $\boldsymbol{R}^+_{m,j}$ simply by exchanging its ith column with its mth one, then the forgetting algorithm can be summarized as follows:

For $j = 1, \cdots, n$

1) obtain $(\boldsymbol{R}^*_{m,j})^+$ from $\boldsymbol{R}^+_{m,j}$
2) obtain $\boldsymbol{P}_j$ and $\boldsymbol{q}_j$ from $(\boldsymbol{R}^*_{m,j})^+$
3) obtain $\boldsymbol{d}_j$
4) compute $(\boldsymbol{R}^*_{m-1,j})^+$
5) compute the unknown vectors $\tilde{\boldsymbol{W}}^T_{m-1,j}$
6) expand $\tilde{\boldsymbol{W}}^T_{m-1,j}$ with zero elements until $\boldsymbol{W}^T_{m-1,j}$ is obtained.

IV. Simulation Results

Simulations consist of three phases. In the first phase, a 3-pattern (6×6)-DTCNN $(r = 1)$ has been synthesized. In the learning phase, an additional pattern has been stored, whereas in the forgetting phase one of the three original patterns has been deleted. In Fig. 1(a) and (b), bipolar images corresponding to the first two phases have been reported. In these figures, noisy images are reported on the left, whereas the desired outputs on the right coincide with the patterns stored in the cellular associative memory. In Fig. 1(c), a forgotten pattern (the second one) is inserted among the proposed inputs. It can be noted that the corresponding output is not an asymptotically stable equilibrium point.

V. Conclusion

In this letter, a new synthesis procedure for associative memories using DTCNN with learning and forgetting capabilities has been presented. These capabilities are provided preserving the cellular architecture and without affecting the existing equilibria.

References

[1] J. Si and A. N. Michel, "Analysis and synthesis of discrete-time neural networks with multilevel threshold functions," in *Proc. 1991 IEEE Int. Symp. Circuits Syst. (ISCAS '91)*, Singapore, vol. 3, pp. 1461–1464.

[2] G. Yen and A. N. Michel, "A learning and forgetting algorithm in associative memories: Results involving pseudo-inverses," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1193–1205, Oct. 1991.

[3] ——, "A learning and forgetting algorithm in associative memories: The eigenstructure method," *IEEE Trans. Circuits Syst., Part II*, vol. 39, pp. 212–225, Apr. 1992.

[4] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1257–1272, Oct. 1988.

[5] H. Harrer, J. A. Nossek, and R. Stelzl, "An analog implementation of discrete-time cellular neural networks," *IEEE Trans. Neural Networks*, vol. 3, no. 3, pp. 466–476, 1992.

[6] D. Liu and A. N. Michel, "Cellular neural networks for associative memories," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 119–121, Feb. 1993.

[7] M. Brucoli, L. Carnimeo, and G. Grassi, "A new synthesis procedure for associative memories using Discrete-Time Cellular Neural Networks," in *1994 Symp. Nonlinear Theory and Its Applicat. (Nolta'94)*, Kagoshima, Japan, Oct. 6–8, 1994.

[8] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications.* New York: Wiley, 1974.

Semi-Implicit Differential-Algebraic Equations Constitute a Normal Form

Gunther Reißig

***Abstract*— Continuously differentiable functions, the total derivative, or a partial derivative of which is of constant rank, play a part in many engineering problems. One usually exploits this property of constancy of rank by applying the Rank Theorem. However, in case only a partial derivative is of constant rank, which is the natural situation for functions involved in Differential-Algebraic Equations (DAE's), this theorem does not apply immediately. In this letter, we generalize known results to the latter case. More precisely, we give a parameterized version of the Rank Theorem and results on functional dependence and present a normal form for a class of nonlinear equations. Although these results are general in nature, the fundamental conclusion with respect to DAE's is that here the normal form exactly corresponds to semi-implicit DAE's. We also generalize results from the solution theory of DAE's in case differential geometric techniques fail to apply. Such DAE's occur, for example, in the analysis of certain circuits.**

I. Introduction

Consider the equation[1]

$$f(x) = 0 \tag{1}$$

where $U \subseteq \mathbb{R}^n$ is open, $f: U \to \mathbb{R}^m \in C^1$, and assume that $\operatorname{rank} Df(x) = \text{const.}$ as long as $x \in U$. Assume further that $f(x_0) = 0$ for some $x_0 \in U$ and let $Q: \mathbb{R}^m \to \mathbb{R}^m$ be a projector onto some algebraic complement of $\operatorname{im} Df(x_0)$. Then, by virtue of the Rank Theorem [1], $Q \circ f$ is functionally dependent on $(\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f$ on some neighborhood of x_0 [2, Problem 4.4.c]. From that, it immediately follows that for x in some neighborhood of x_0, (1) has exactly the same solutions as

$$(\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f(x) = 0.$$

In particular, if $f(x) = (f_1(x), \cdots, f_m(x))$ and $\operatorname{rank} Df(x_0) = r$, this means that (1) has exactly the same solutions as

$$\begin{aligned} f_{\kappa(1)}(x) &= 0 \\ &\vdots \\ f_{\kappa(r)}(x) &= 0 \end{aligned} \tag{2}$$

in some neighborhood of x_0, where $\kappa: \{1, \cdots, m\} \to \{1, \cdots, m\}$ is a suitable bijection.

We now ask for generalizations of these results to parameter dependent equations, i.e., equations of form

$$f(p, x) = 0 \tag{3}$$

where $P \subseteq \mathbb{R}^k$ and $U \subseteq \mathbb{R}^n$ are open, $(p_0, x_0) \in P \times U$, $f: P \times U \to \mathbb{R}^m \in C^1$, and $\operatorname{rank} D_2 f(p, x) = \text{const.}$ as long as $(p, x) \in P \times U$.

Manuscript received Decenber 14, 1994; revised February 24, 1995. This paper was recommended by Associate Editor H.-D. Chiang.

The author is with the Institut für Grundlagen der Elektotechnik und Elektronik, Technische Universität Dresden, Germany.

IEEE Log Number 9412333.

[1] We use the following notation: $Df(x)$ is the derivative of f at x, while $D_i f$ denotes the partial derivative of f with respect to its ith argument. $\operatorname{im} f$ denotes the image of f, $f|_A$ denotes the restriction of f to the set A, and id is the identity mapping.

Clearly, for any $p \in P$, the Rank Theorem applies to $f(p, \cdot)$ and one obtains

$$\Phi(p, \cdot) \circ f(p, \cdot) \circ \Psi(p, \cdot) = D_2 f(p_0, x_0)|_{\mathrm{dom}\Psi(p, \cdot)} \quad (4)$$

where, for each p, the mappings $\Phi(p, \cdot)$, and $\Psi(p, \cdot)$ are diffeomorphisms defined on suitable neighborhoods of $f(p, x_0)$ and 0, respectively [3].

There are two difficulties with this approach.

1) It is not at all clear whether Φ and Ψ can be chosen in such a way that they are defined on neighborhoods of $(p_0, f(p_0, x_0))$ and $(p_0, 0)$, respectively.
2) If Φ and Ψ are defined on open sets, the Rank Theorem does not provide any information on continuity or differentiability properties of these mappings with respect to their first argument, i.e., with respect to the parameter.

In this note, we give a Corollary to the Rank Theorem in [1], which states that the mappings Φ and Ψ in (4) can be chosen to be defined on open sets and also provides information on differentiability of these functions with respect to their first argument.

We then consider, in complete analogy to the parameter independent case, the relation between constancy of rank of a partial derivative and functional dependence and give a normal form for the nonlinear equation (3).

In many problems, such as the analysis of electrical circuits [4]–[6] and mechanical systems [5], [6], modeling of chemical reactions [5], and numerical solution of partial differential equations by discretization [5], one has to deal with equations of form

$$0 = f(x, \dot{x}). \quad (5)$$

These are usually referred to as Differential-Algebraic Equations (DAE's), if $f: V \to \mathbb{R}^n \in C^1$, $V \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is open, and $\operatorname{rank} D_2 f(x, y) = \text{const.}$ as long as $(x, y) \in V$ [6]. DAE's arising in one of the fields mentioned above are of a special form each. However, in order to establish a uniform theory that covers all of the above applications, it is important to treat the general case (5). This can be considerably simplified by using the normal form of (3), which, in case of DAE's, turns out to exactly correspond to semi-implicit DAE's. Further, we generalize known results from the solution theory of DAE's by applying the latter fact.

II. A Corollary to the Rank Theorem

We first give the classical result [1, 10.3.1.].[2]

II.1. Theorem (Rank Theorem): Let $U \subseteq \mathbb{R}^n$ be open, $x_0 \in U$, $f: U \to \mathbb{R}^m \in C^s$ ($s \in \mathbb{N} \cup \{\infty, \omega\}$), and $\operatorname{rank} Df(x) = \text{const.}$ for $x \in U$.

Then there are neighborhoods $W \subseteq \mathbb{R}^n$, $W' \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$, and $V' \subseteq \mathbb{R}^m$ of 0, x_0, $f(x_0)$, and 0, respectively, and C^s-diffeomorphisms $\Psi: W \to W'$ and $\Phi: V \to V'$, that

$$\Phi \circ f \circ \Psi = Df(x_0)|_W.$$

□

We now give a parameterized version of the Rank Theorem.

II.2. Corollary: Let $P \subseteq \mathbb{R}^k$ and $U \subseteq \mathbb{R}^n$ be open, $(p_0, x_0) \in P \times U$, $f: P \times U \to \mathbb{R}^m \in C^s$ ($s \in \mathbb{N} \cup \{\infty, \omega\}$), and $\operatorname{rank} D_2 f(p, x) = \text{const.}$ for $(p, x) \in P \times U$.

Then there are neighborhoods $\tilde{P} \subseteq \mathbb{R}^k$, $W \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ of p_0, 0, and $f(p_0, x_0)$, respectively, and C^s-mappings $\Psi: \tilde{P} \times W \to \mathbb{R}^n$ and $\Phi: \tilde{P} \times V \to \mathbb{R}^m$ that for any $p \in \tilde{P}$ the following hold:

i) $\Phi(p, \cdot) \circ f(p, \cdot) \circ \Psi(p, \cdot) = D_2 f(p_0, x_0)|_W$.
ii) $\Phi(p, \cdot): V \to \Phi(p, V)$ and $\Psi(p, \cdot): W \to \Psi(p, W)$ are C^s-diffeomorphisms.

Further, the mappings Φ and Ψ have the following properties:

iii) $\Phi(p_0, f(p_0, x_0)) = 0$, $D_1\Phi(p_0, f(p_0, x_0)) = -D_1 f(p_0, x_0)$, and $D_2\Phi(p_0, f(p_0, x_0)) = \mathrm{id}|_{\mathbb{R}^m}$.
iv) $\Psi(p_0, 0) = x_0$, $D_1\Psi(p_0, 0) = 0$, and $D_2\Psi(p_0, 0) = \mathrm{id}|_{\mathbb{R}^n}$. □

Proof: The crucial trick of this proof is to consider the following mapping:

$$F: P \times U \to \mathbb{R}^k \times \mathbb{R}^m: (p, x) \mapsto (p, f(p, x)).$$

Obviously, F is C^s if f is, and, for $h \in \mathbb{R}^k$, $l \in \mathbb{R}^n$, we have $DF(p, x)(h, l) = (h, Df(p, x)(h, l))$, and therefore, $\ker DF(p, x) = \{0\} \times \ker D_2 f(p, x)$. Hence, $DF(p, x)$ is of constant rank for $(p, x) \in P \times U$, since $D_2 f(p, x)$ is. Further considerations, conclusions from Rank Theorem and Implicit Function Theorem (IFT) in the main, yield the assertions [7]. □

III. Applications

A. Functional Dependence and a Normal Form for a Class of Nonlinear Equations

In this section, we will investigate functional dependence between parts of those functions that Corollary II.2. applies to. A basic result is the following:

III.1. Corollary (Functional Dependence): Let P, U, (p_0, x_0), and f as in Corollary II.2. Let further F_2 be some algebraic complement of $\operatorname{im} D_2 f(p_0, x_0)$, $\pi: \mathbb{R}^m \to \mathbb{R}^m$ be the projector onto F_2 along $\operatorname{im} D_2 f(p_0, x_0)$, and $Q: \mathbb{R}^m \to \mathbb{R}^m$ be a projector onto F_2. Then there are neighborhoods $\tilde{P} \subseteq \mathbb{R}^k$, $\tilde{U} \subseteq \mathbb{R}^n$, and $V \subseteq \ker Q$ of p_0, x_0, and $(\mathrm{id}|_{\mathbb{R}^m} - Q) f(p_0, x_0)$, respectively, and some C^s-mapping $H: \tilde{P} \times V \to F_2$ that the following hold:

i) $\forall_{(p, x) \in \tilde{P} \times \tilde{U}}\, Q \circ f(p, x) = H(p, (\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f(p, x))$.
ii) $D_1 H(p_0, (\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f(p_0, x_0)) = \mathrm{id}|_{F_2} \circ \pi \circ D_1 f(p_0, x_0)$.
iii) $D_2 H(p_0, (\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f(p_0, x_0)) = -\mathrm{id}|_{F_2} \circ \pi|_{\ker Q}$. □

Proof: By Corollary II.2., we have

$$\Phi(p, \cdot) \circ f(p, \cdot) \circ \Psi(p, \cdot)(x) = D_2 f(p_0, x_0)x$$

as long as (p, x) in some neighborhood of $(p_0, 0)$. Conclude from the IFT [2] that there is some mapping $\hat{\Psi}$ that

$$\Phi(p, \cdot) \circ f(p, \cdot)(x) = D_2 f(p_0, x_0)\hat{\Psi}(p, x)$$

for (p, x) in some neighborhood of (p_0, x_0) [7]. We then have

$$\pi \circ \Phi(p, f(p, x)) = 0$$

as long as (p, x) in some neighborhood of (p_0, x_0). Applying the IFT to the mapping $g: (p, x_1, x_2) \mapsto \mathrm{id}|_{F_2} \circ \pi \circ \Phi(p, x_1 + x_2)$, for (p, x_1, x_2) in a suitable open set, completes the proof [7]. □

We are now interested in the set of solutions of (3), for f as in Corollaries II.2. and III.1.

In case f is independent of p, which is the classical situation where the Rank Theorem can be applied, a certain part of (3) can be completely neglected, namely, (1) is equivalent to (2), see Section I. For the parameter dependent case, the analogous result is as follows:

[2] $\mathbb{N}$ denotes the set of natural numbers 1, 2, $\cdots$, the set of s times continuously differentiable functions from E into F is denoted by $C^s(E, F)$, for $s \in \mathbb{N} \cup \infty$, and the set of analytic functions from E into F is denoted by C^ω. We will omit naming E and F explicitly unless these abbreviations lead to misunderstandings.

III.2. Corollary (Normal Form): Let f be as in Corollaries II.2. and III.1., consider (3) and assume that $f(p_0, x_0) = 0$. Let further be F_2, π, and Q as in Corollary III.1.

Then there are neighborhoods $\tilde{P} \subseteq \mathbb{R}^k$ and $\tilde{U} \subseteq \mathbb{R}^n$ of p_0 and x_0, respectively, and a C^s-mapping $g: \tilde{P} \to F_2$ that the system of equations

$$0 = (\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f(p, x)$$
$$0 = g(p)$$

has exactly the same solutions as (3) in $\tilde{P} \times \tilde{U}$ and $Dg(p_0) = \mathrm{id}|_{F_2} \circ \pi \circ D_1 f(p_0, x_0)$. □

Proof: Apply Corollary III.1., set $g := H(\cdot, 0)$, and we are done [7]. □

III.3. Remark:

i) In [8], there is a result similar to Corollary III.2. However, we do not assume that $f^{-1}(0)$ is a manifold, which is an essential hypothesis in [8]. Indeed, circuits containing multipliers or hysteresis elements lead to (3) and (5) with $f^{-1}(0)$ being not a manifold [9].
ii) Corollary II.2. is an independent result that led to the normal form of (3) given in Corollary III.2. However, one can obtain this normal form without using any Rank Theorem. I thank the reviewer who pointed this out to me. □

III.4. Remark:

i) In case $f(p, x) = (f_1(p, x), \cdots, f_m(p, x))$ and $r = \mathrm{rank}\, D_2 f(p_0, x_0)$ we can choose F_2 and Q such that (3) has the same solutions in $\tilde{P} \times \tilde{U}$ as

$$0 = f_{\kappa(1)}(p, x)$$
$$\vdots$$
$$0 = f_{\kappa(r)}(p, x)$$
$$0 = g_1(p)$$
$$\vdots$$
$$0 = g_{m-r}(p),$$

where $\kappa: \{1, \cdots, m\} \to \{1, \cdots, m\}$ is a suitable bijection and $g(p) = (g_1(p), \cdots, g_{m-r}(p))$.
ii) $Df(p_0, x_0)$ is surjective if and only if $Dg(p_0)$ is.
iii) Note that $\mathrm{im}\, D_2((\mathrm{id}|_{\mathbb{R}^m} - Q) \circ f)(p_0, x_0) = \ker Q$, and hence, the structure of the zero set of $f(p, \cdot)$ is known from the Surjective Implicit Function Theorem [2] if the parameter p is sufficiently close to p_0 and $g(p) = 0$.
iv) $f^{-1}(0)$ is a C^q-submanifold of $\mathbb{R}^k \times \mathbb{R}^n$ near (p_0, x_0) if $g^{-1}(0)$ is a C^q-submanifold of $\mathbb{R}^k$ near p_0, for $1 \leq q \leq s$. □

A proof to the foregoing Remark III.4. can be found in [7].

B. Application to Differential-Algebraic Equations

The special case that shall be of primary interest here, is the semi-implicit case of (5), namely, we call (5) a *semi-implicit DAE* [6] if $V = V_x \times V_y$ and there are functions $f_1: V \to \mathbb{R}^r$ and $f_2: V_x \to \mathbb{R}^{n-r}$ that

$$f(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x) \end{pmatrix} \quad \text{and} \quad D_2 f_1(x, y) \text{ is surjective}$$

as long as $(x, y) \in V$. The fundamental result of this section is that the class of semi-implicit DAE's even is a normal form.

III.5. Corollary: Consider DAE (5) and let $(x_0, y_0) \in f^{-1}(0)$. Then there is some semi-implicit DAE that has exactly the same C^1-solutions as DAE (5) locally near (x_0, y_0).

More precisely, if $f \in C^s(s \in \mathbb{N} \cup \{\infty, \omega\})$, and $r = \mathrm{rank}\, D_2 f(x_0, y_0)$, then there are neighborhoods V_x of x_0 and V_y of y_0, $V_x \times V_y \subseteq V$, and C^s-functions $f_1: V \to \mathbb{R}^r$ and $f_2: V_x \to \mathbb{R}^{n-r}$ that

$$0 = f_1(x, \dot{x})$$
$$0 = f_2(x)$$

has exactly the same C^1-solutions as

$$0 = f|_{V_x \times V_y}(x, \dot{x}).$$

Moreover, f_1 and f_2 can be chosen to have the following properties:

i) $f_1 = \Pi_1 \circ (\mathrm{id}|_{\mathbb{R}^n} - Q) \circ f|_{V_x \times V_y}$
ii) $f_2(x_0) = 0$
iii) $Df_2(x_0) = \Pi_2 \circ \pi \circ D_1 f(x_0, y_0)$,

where $\{e_i | \; i \in \{1, \cdots, n\}\}$ is the canonical basis of $\mathbb{R}^n$, $\kappa: \{1, \cdots, n\} \to \{1, \cdots, n\}$ is a suitable bijection and $F_2 = \mathrm{span}\{e_{\kappa(1)}, \cdots, e_{\kappa(n-r)}\}$ is an algebraic complement of $\mathrm{im}\, D_2 f(x_0, y_0)$, $Q: \mathbb{R}^n \to \mathbb{R}^n$ is the orthogonal projector onto F_2, $\pi: \mathbb{R}^n \to \mathbb{R}^n$ is the projector onto F_2 along $\mathrm{im}\, D_2 f(x_0, y_0)$, and $\Pi_1: \ker Q \to \mathbb{R}^r$ and $\Pi_2: F_2 \to \mathbb{R}^{n-r}$ are linear isomorphisms. □

Proof: Apply Corollary III.2. The special choice of F_2 is practicable by STEINITZ's Exchange Lemma, and $D_2 f_1(x, y)$ is surjective for all $(x, y) \in V_x \times V_y$ by Remark III.4.iii). □

When dealing with problems of existence and uniqueness of solutions of DAE (5), one is necessarily led to the following problem:

Let $x: I \to \mathbb{R}^n \in C^1$ be a solution to DAE (5) and let $t_0 \in I$. Under which conditions to f does the relation

$$D_1 f(x(t_0), \dot{x}(t_0))\dot{x}(t_0) \in \mathrm{im}\, D_2 f(x(t_0), \dot{x}(t_0)) \qquad (6)$$

hold?

In [10, Lemma 2.1.] it is proved that relation (6) holds if $Df(x(t_0), \dot{x}(t_0))$ is surjective. However, from Corollary III.5., it immediately follows that this restriction to cases where $f^{-1}(0)$ is a manifold is not necessary.

III.6. Corollary: Let $x: I \to \mathbb{R}^n \in C^1$ be a solution to DAE (5) and let $t_0 \in I$. Then relation (6) holds. □

Proof: For semi-implicit DAE's, the assertion is trivial. Applying Corollary III.5. completes the proof [7]. □

IV. Conclusion

We gave a Corollary to the Rank Theorem that generalizes the classical result to the case when a partial derivative of a function is of constant rank. We then investigated the relation between functional dependence of certain parts of a function and constancy of rank of a partial derivative of that function.

These results were used to establish a normal form for a class of nonlinear equations. In case of Differential-Algebraic Equations (DAE's) it turned out that this normal form exactly corresponds to semi-implicit DAE's. This is a cornerstone of a solution theory for DAE's (5) in case $f^{-1}(0)$ is not a manifold. Such DAE's arise in the analysis of even simple circuits if these, for example, contain multipliers or hysteresis elements [9].

Let us finally remark that our assumption on constancy of rank of $D_2 f$ is essential. There are simple implicit Differential Equations that do not satisfy it, and our theorems do not apply in this case.

References

[1] J. Dieudonné, *Foundations of Modern Analysis*, vol. 1, 7th ed. New York and London: Academic, 1968.

[2] E. Zeidler, "Fixed-point theorems," in *Nonlinear Functional Analysis and its Applications*, vol. 1 (corrected ed.) New York, Berlin: Springer-Verlag 1993.

[3] E. Griepentrog, "Index reduction methods for differential-algebraic equations," in *Berlin Seminar on Differential-Algebraic Equations, Seminar Bericht Nr. 92-1*, E. Griepentrog, M. Hanke, and R. März, Eds. Berlin, Germany: Fachbereich Mathematik der HU Berlin, Jan. 1992, pp. 14–29.

[4] B. C. Haggman and P. R. Bryant, "Solutions of singular constrained differential equations: A generalization of circuits containing capacitor-only loops and inductor-only cutsets," *IEEE Trans. Circuits Syst.*, vol. 31, pp. 1015–1029, Dec. 1984.

[5] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations.* New York, Amsterdam, London: North-Holland, 1989.

[6] W. C. Rheinboldt, "Chapter 6: The theory and numerics of differential-algebraic equations," in *Advances in Numerical Analysis*, vol. I. Lancaster, UK, 1990, "Nonlinear partial differential equations and dynamical systems," *Proc. 4th. Summer School*, pp. 237–275.

[7] G. Reißig, "Semi-implicit differential-algebraic equations constitute a normal form," TU Preprint ET-IEE-18-1994, TU Dresden, Oct. 12, 1994.

[8] P. J. Rabier and W. C. Rheinboldt, "A geometric treatment of implicit differential-algebraic equations," *J. Differential Equations*, vol. 109, pp. 110–146, 1994.

[9] R. W. Newcomb and G. Reißig, private communication, 1994.

[10] P. J. Rabier and W. C. Rheinboldt, "A general existence and uniqueness theory for implicit differential-algebraic equations," *Differential and Integral Equations*, vol. 4, no. 3, pp. 563–582, May 1991.

On Approximation of Linear Functionals on L_p Spaces

Irwin W. Sandberg and Ajit Dingankar

***Abstract*— In a recent paper certain approximations to continuous nonlinear functionals defined on an L_p space $(1 < p < \infty)$ are shown to exist. These approximations may be realized by sigmoidal neural networks employing a linear input layer that implements finite sums of integrals of a certain type. In another recent paper similar approximation results are obtained using elements of a general class of continuous linear functionals. In this note we describe a connection between these results by showing that every continuous linear functional on a compact subset of L_p may be approximated uniformly by certain finite sums of integrals.**

I. Introduction

One of the earliest results in the area of neural networks is the proposition that any continuous real function defined on a compact subset of $\mathbb{R}^k$ (k an arbitrary positive integer) can be approximated arbitrarily well using a single-hidden-layer network with sigmoidal nonlinearities (see, for example, [2]). Among other results in the literature concerning arbitrarily good approximation that concern more general types of "target" functionals, different network structures, other nonlinearities, and various measures of approximation errors is the proposition in [5], [4] that any continuous real nonlinear functional on a compact subset of a real normed linear space can be approximated arbitrarily well using a single-hidden-layer neural network with a linear functional input layer and exponential (or polynomial or sigmoidal) nonlinearities.

In an interesting paper [1] by Chen and Chen similar results concerning single-hidden-layer neural networks with sigmoidal nonlinearities are proved for continuous real functionals on a compact subset of an L_p space $(1 < p < \infty)$. One of the main results in [1] is the proposition that the linear functionals in the input layer may have the special form of a finite sum of integrals of a certain type. Here we connect these results by showing that every continuous linear functional can be approximated arbitrarily well on compact sets by finite sums of the type used in [1].

More specifically, the main result in [4] shows that any real continuous functional defined over a compact subset of a real normed linear space can be approximated arbitrarily well with a neural network employing an input layer of functionals that can be taken to be linear, and a single-hidden-layer implementing nonlinearities that, for example, can be taken to be sigmoidal. To describe the result proved there, let X be a real normed linear space, and let X^* be the set of bounded linear functionals on X (i.e., the set of bounded linear maps from X to the reals $\mathbb{R}$). Given a compact subset C of X, let Y be any set of continuous maps from X to $\mathbb{R}$ that is dense in X^* on C, in the sense that for each $\phi \in X^*$ and any $\rho > 0$ there is a $y \in Y$ such that $|\phi(x) - y(x)| < \rho$, $x \in C$. Also, let U be any set of continuous maps $u : \mathbb{R} \to \mathbb{R}$ such that given $\sigma > 0$ and any bounded interval $(\beta_1, \beta_2) \subset \mathbb{R}$ there exists a finite number of elements $u_1, \ldots, u_\ell$ of U for which $|\exp(\beta) - \sum_j u_j(\beta)| < \sigma$ for $\beta \in (\beta_1, \beta_2)$.[1] The result in [4] is this: Let g be a real-valued continuous map defined on C. Then given $\epsilon > 0$ there are a positive integer k, real numbers $c_1, \ldots, c_k$, elements $u_1, \ldots, u_k$ of U, and elements $y_1, \ldots, y_k$ of Y such that

$$|g(x) - \sum_j c_j u_j[y_j(x)]| < \epsilon$$

for $x \in C$.

The proof in [4][2] of the result described above shows that the result can be slightly sharpened in that "$y_1, \cdots, y_k$ of Y" can be replaced with "$y_1, \cdots, y_k$ of Y_ρ for some $\rho > 0$," where each Y_ρ is any set of continuous maps from X to $\mathbb{R}$ with the property that given $\phi \in X^*$ with $\|\phi\| \le 1$ there is a $y \in Y_\rho$ such that $|\phi(x) - Y(x)| < \rho$, $x \in C$. Now suppose that $X = L_p$ with $1 < p < \infty$, and that $U = \{u : u(\beta) = cs(w\beta + \gamma),\ c, w, \gamma \in \mathbb{R}\}$, where s is a sigmoidal function. In this setting we obtain one of the main results in [1] (assuming continuous as opposed to generalized [1] sigmoids) as a special case of the above slightly-sharpened result, since our theorem below and a simple observation in Section 2.3 about its proof permit us to take the Y_ρ to be sets of finite linear combinations of integral functionals of the form considered in [1].[3] That is, in the L_p setting, for each ρ there is an $h > 0$ such that we can take the value at x of the elements y of Y_ρ to be given by certain

[1] Of course we can take U to be the set whose only element is $\exp(\cdot)$, or the set $\{u : u(\beta) = (\beta)^n/n!,\ n \in \{0, 1, \ldots\}\}$.

[2] We take this opportunity to correct three typos in the proof in [4], all of which occur in the second column of page 373: The third β_1 on line 2 should be replaced with β_2, and $|\sum_j d_j \exp[z_j(x)] - \sum_j d_j \exp[y_j(x)]|$ and $(2\epsilon)/3$ should be added in the obvious places on lines 7 and 17, respectively.

[3] We also extend the result in [1] to $p = 1$.

Manuscript received October 27, 1994; revised February 23, 1995. This paper was recommended by Associate Editor H.-D. Chiang.

The authors are with the Department of Electrical and Computer Engineering, University of Texas, Austin, TX 78712 USA.

IEEE Log Number 9412332.