Eighth Mississippi State - UAB Conference on Differential Equations and Computational Simulations. *Electronic Journal of Differential Equations*, Conf. 19 (2010), pp. 31–36.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

MODELS OF LEARNING AND THE POLAR DECOMPOSITION OF BOUNDED LINEAR OPERATORS

FERNANDA BOTELHO, ANNITA DAVIS

Abstract. We study systems of differential equations in $\mathcal{B}(\mathcal{H})$, the space of all bounded linear operators on a separable complex Hilbert space $\mathcal{H}$ equipped with the operator norm. These systems are infinite dimensional generalizations of mathematical models of learning. We use the polar decomposition of operators to find an explicit form for solutions. We also discuss the standard questions of existence and uniqueness of local and global solutions, as well as their long-term behavior.

1. Introduction

Learning models are devices that reproduce features of the human's ability to interact with the environment. Such models are typically implemented in a feed-forward neural network consisting of a finite number of interconnected units, designated neurons. Each neuron has the capability of receiving, combining and processing quantifiable information. The interconnection among neurons is represented by a net of channels through which information flows. This translates the activity of a brain at the synaptic level. It is natural to predict that information changes in this system. This is represented by the action of multiplicative factors, designated connecting weights. A mathematical model of learning should encompass an algebraic interpretation of this phenomenon, in general, given as a system of differential equations. The stability of a learning model is of crucial importance since it provides information on how the device emerges after exposed to a set of initial conditions. If stability occurs, the result yields a set of weight values that represents learning after the exposure to an initial stimulus. In this paper we address this aspect of learning theory, often referred in literature as "unsupervised learning". Several researchers have proposed systems that perform unsupervised learning, see [3, 13, 16]. In [17] Oja introduced a learning model that behaves as an information filter with the capability to adapt from an internal assignment of initial weights. This approach uses the principal component analyzer statistical method to perform a selection of relevant information. More recently, Adams [2] proposed a generalization of Oja's model by incorporating a probabilistic parameter. Such

2000 *Mathematics Subject Classification.* 34G20, 47J25.
Key words and phrases. Nonlinear systems; learning models;
polar decomposition of operators.

Published September 25, 2010.

a parameter captures the possible creation of temporary synapses or channels in an active network. We refer the reader to [5] for a detailed interpretation of the Cox-Adams model for a network with n input neurons and a single output neuron. This model is given by the system of differential equations:

$$\frac{dW}{dt} = TCW - WW^TCW$$

with W a column of connecting weights and C is a symmetric matrix. Each entry value in C is equal to the expected input correlation between two neurons. The matrix T is a non-singular tri-diagonal matrix $[t_{ij}]_{i,j}$ given by

$$t_{ij} = \begin{cases} 1-\varepsilon & \text{if } i = j \\ \varepsilon/2 & \text{if } |i-j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

This matrix translates the synaptic formation according to some probability ϵ.

In this paper we consider a generalization of this system to an infinite dimensional setting. This better reflects the complexity of real systems where continuous activity occurs. Our new setting is the Banach space $\mathcal{B}(\mathcal{H})$ of bounded operators on a separable Hilbert space . We consider the system

$$\frac{dZ}{dt} = TMZ - ZZ^*MZ, \tag{1.1}$$

with T representing an invertible, positive, self-adjoint operator on $\mathcal{H}$ and M a self-adjoint operator on $\mathcal{H}$. Particularly interesting examples are the tridiagonal self-adjoint operators, see [7, 8, 10, 11].

The operator valued, time dependent Z now represents the continuous change of connecting weights according to the rule described in equation (1.1). We present a scheme that explicitly solves system (1.1). First a natural change of variables reduces (1.1) to a static system where no synaptic formation occurs. However, the probabilistic effect transfers to the input correlation operator M. System (1.1) reduces to an Oja type model. We follow a strategy employed in [4]. The main tool is the polar decomposition of operators that allows us to derive a scalar system and a polar system associated with the original system. Both systems are solved explicitly. These two solutions combined define the local solution for the original system, given certain mild constrains on the initial conditions. The explicit form for local solutions is now used to derive the existence of global solutions and for the stability analysis.

2. Background Results

In this section we summarize the techniques used in [4] to solve the generalization of Oja-Karhunen's model on a separable Hilbert space. We recall that the generalized Oja-Karhunen model is given as follows

$$\begin{gathered} \dot{Z} = MZ - ZZ^*MZ \\ Z(0) = Z_0. \end{gathered} \tag{2.1}$$

The time dependent variable Z has values in $\mathcal{B}(\mathcal{H})$. The operator Z^* is the adjoint of Z and M is a normal operator on $\mathcal{H}$.

Classical fixed point theorems allow us to assure the local existence and uniqueness of solutions for system (2.1), see [15, p. 405].

Theorem 2.1 ([4])**.** *Let Z_0 be a bounded operator in $\mathcal{B}(\mathcal{H})$. If $F : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is a Lipschitz function then there exists a positive number ε and a unique differentiable map $Z : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that $\dot{Z} = F(Z)$ and $Z(0) = Z_0$.*

It is straightforward to check that $Z \to MZ - ZZ^*MZ$ is a Lipschitz function and hence local existence and uniqueness of solutions of (2.1) follow from Theorem 2.1.

The technique employed in the solution scheme uses the well known polar decomposition of bounded operators, see Ringrose [19]. A bounded operator Z can be written as the product of a partial isometry P and a hermitian operator $\sqrt{ZZ^*}$. The operator $\sqrt{ZZ^*}$ represents the unique positive square root of ZZ^* and the operator P satisfies the equation $PP^*P = P$. The bounded operator Z is decomposed as follows: $Z = \sqrt{ZZ^*}P$. This decomposition is unique and is called the polar decomposition of the operator Z.

In [4] we applied the polar decomposition to construct two new systems associated with (2.1). The solutions of these systems are the scalar and the polar components of the solution of the original system.

For $t \in (-\varepsilon, \varepsilon)$ we denote by $Z(t)$ a local solution of (2.1) and we set $V(t) = Z(t)Z(t)^*$. It is a straightforward calculation to verify that $V(t)$ is a local solution of the system

$$\begin{gathered} \dot{V} = MV + VM^* - VMV - VM^*V \\ V(0) = Z_0 Z_0^*. \end{gathered} \tag{2.2}$$

If M and Z_0 commute, Fuglede-Putman Theorem (cf. Furuta [12, p. 67]) and Piccard's iterative method (Hartman [15, pp. 62-66]) imply that the family $\{V(t)\}_{t\in(-\epsilon,\epsilon)}$ defines a path of positive operators that commute with M and M^*. Furthermore $\{V(t)\}_{t\in(-\epsilon,\epsilon)}$ is a family of commuting operators. Thus system (2.2) can be written as

$$\dot{V} = (M + M^*)(V - V^2), \quad \text{with } V(0) = V_0.$$

Since Z_0 is an invertible operator in $\mathcal{H}$, for some $\varepsilon > 0$, $V(t)$ is also an invertible operator for $t \in (\varepsilon, \varepsilon)$, cf [9]. We have that

$$\frac{d}{dt}(V^{-1}) = -V^{-2}\dot{V}.$$

Using the commutativity of M and V, we found that

$$\frac{d}{dt}(V^{-1}) = (M + M^*) - (M + M^*)V^{-1},$$

which is a first-order linear differential equation, see [1]. A generalization of standard techniques of integrating factors, appropriately generalized to infinite dimensions (see [4] pg 101), imply that $V(t) = [I + (V_0^{-1} - I)exp(-(M + M^*)t)]^{-1}$ for $t \in (-\epsilon, \epsilon)$.

We derive the polar system associated with (2.1). This is a first order non autonomous linear differential equation. For simplicity of notation we set $V^{1/2} = \sqrt{ZZ^*}$. Using the commutativity of V and M, we obtain

$$\begin{gathered} \dot{P} = -\frac{1}{2}(M - M^*)(V(t) - I)P \\ P(0) = P_0. \end{gathered} \tag{2.3}$$

Properties of exponential operator-valued functions imply that

$$P(t) = \exp\Big(\int_0^t -\frac{1}{2}(M-M^*)(V(\xi)-I)d\xi\Big)P_0 \quad \text{for } |t|<\varepsilon)$$

is a solution of the polar system (2.3). These considerations are now summarized in the following theorem.

Theorem 2.2 ([4]). *If Z_0 is invertible and commutes with the normal operator M, then there exist $\varepsilon > 0$ and a unique differentiable mapping $Z : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that*

$$\begin{gathered} \dot{Z} = MZ - ZZ^*MZ \\ Z(0) = Z_0, \end{gathered} \tag{2.4}$$

if and only if $Z(t) = V(t)^{1/2}P(t)$,

$$V(t) = [I + (V_0^{-1} - I)\exp(-(M+M^*)t)]^{-1},$$

and

$$P(t) = \exp\Big(\int_0^t -\frac{1}{2}(M-M^*)(V(\xi)-I)d\xi\Big)P_0.$$

3. General Solution for the Cox-Adams Learning Model

We recall that the Cox-Adams learning model is

$$\begin{gathered} \frac{dZ}{dt} = TMZ - ZZ^*MZ \\ Z(0) = Z_0 \end{gathered} \tag{3.1}$$

with T representing an invertible, positive, self-adjoint operator, and M self-adjoint on $\mathcal{H}$. Theorem 2.1 implies the local existence and uniqueness of solutions.

Since T is positive and invertible, we rewrite equation (3.1) as follows

$$\frac{dZ}{dt} = (\sqrt{T}\sqrt{T})M\big(\sqrt{T}(\sqrt{T})^{-1}\big)Z - ZZ^*\big((\sqrt{T})^{-1}\sqrt{T}\big)MZ,$$

equivalently,

$$(\sqrt{T})^{-1}\frac{dZ}{dt} = \sqrt{T}M\sqrt{T}(\sqrt{T})^{-1}Z - (\sqrt{T})^{-1}ZZ^*(\sqrt{T})^{-1}\sqrt{T}MZ.$$

We set $W = (\sqrt{T})^{-1}Z$ and $S = \sqrt{T}M\sqrt{T}$. We observe that S is a hermitian operator. Then system (3.1) becomes

$$\begin{gathered} \dot{W} = SW - WW^*SW \\ W(0) = W_0, \end{gathered}$$

where $W_0 = (\sqrt{T})^{-1}Z_0$.

Proposition 3.1. *If W_0 is invertible and commutes with the hermitian operator S, then there exist $\varepsilon > 0$ and a unique differentiable mapping $W : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that*

$$\dot{W} = SW - WW^*SW \quad \textit{and} \quad W(0) = W_0, \tag{3.2}$$

if and only if $W(t) = V(t)^{1/2}P(t)$, with

$$V(t) = [I + (V_0^{-1} - I)\exp(-2St)]^{-1} \tag{3.3}$$

$$P(t) = P_0. \tag{3.4}$$

Since the operator S is hermitian ($S = S^*$), the proof of the above lemma follows from Theorem 2.2

Theorem 3.2. *Let T be an invertible, positive, self-adjoint operator and M a hermitian operator. If Z_0 is invertible and commutes with M, then there exist $\varepsilon > 0$ and a unique differentiable mapping $Z : (-\varepsilon, \varepsilon) \to \mathcal{B}(\mathcal{H})$ such that*

$$\begin{gathered} \dot{Z} = TMZ - ZZ^*MZ \\ Z(0) = Z_0, \end{gathered} \tag{3.5}$$

if and only if $Z(t) = (TV(t))^{1/2}P(t)$,

$$V(t) = \big[I + \big(\sqrt{T}(Z_0Z_0^*)^{-1}\sqrt{T} - I\big)\exp(-2\sqrt{T}M\sqrt{T}t)\big]^{-1}$$

and $P(t) = P_0$.

The proof of the above theorem follows from Proposition 3.1. The following lemma is used in the stability analysis of the Cox-Adams model.

Lemma 3.3 ([4])**.** *If Z_0 is an invertible operator in $\mathcal{B}(\mathcal{H})$, M is a normal operator that commutes with Z_0, $\|(Z_0Z_0^*)^{-1} - I\| < 1$, and the spectrum of M is strictly positive, then there exists $\epsilon > 0$ so that*

$$I + [(Z_0Z_0^*)^{-1} - I]\exp(-(M + M^*)t)$$

is invertible on the interval $(-\varepsilon, \infty)$ and

$$\lim_{t\to\infty}[I + ((Z_0Z_0^*)^{-1} - I)\exp(-(M + M^*)t)] = I.$$

As a result we have the following corollary.

Corollary 3.4. *Let T be an invertible, positive, self-adjoint operator. If Z_0 is an invertible operator in $\mathcal{B}(\mathcal{H})$, M is a self-adjoint operator that commutes with Z_0, $\|\sqrt{T}(Z_0Z_0^*)^{-1}\sqrt{T} - I\| < 1$, and the spectrum of M is strictly positive, then there exists $\epsilon > 0$ so that*

$$I + [\sqrt{T}(Z_0Z_0^*)^{-1}\sqrt{T} - I]\exp\big(-2\sqrt{T}M\sqrt{T})t\big)$$

is invertible on the interval $(-\varepsilon, \infty)$ and

$$\lim_{t\to\infty}\big[I + [\sqrt{T}(Z_0Z_0^*)^{-1}\sqrt{T} - I]exp\big(-2\sqrt{T}M\sqrt{T}t\big)\big] = I.$$

Remark 3.5. We observe that, under the assumptions in Corollary 3.4, we have

$$\lim_{t\to\infty} Z(t) = P_0.$$

This provides a filtering procedure that selects the polar component of the initial condition.

References

[1] M. Abell, J. Braselton; *Modern Differential Equations*, 2nd ed., Harcourt College Publishers, (2001), 36-421, 197-205.

[2] P. Adams, Hebb, Darwin; *Journal of Theoretical Biology*, **195** (1998), 419-438.

[3] S. Amari; *Mathematical Theory of Neural Networks*, Sangyo-Tosho, Tokyo, 1998.

[4] F. Botelho, A. Davis; *Differential systems on spaces of bounded linear operators, Intl. J. of Pure and Applied Mathematics*, **53** (2009), 95-107.

[5] F. Botelho, J. Jamison; *Qualitative Behavior of Differential Equations Associated with Artificial Neural Networks.* Journal of Dynamics and Differential Equations **16** (2004), 179-204.

[6] J. Conway; *A Course in Functional Analysis*, Graduate Texts in Mathematics, Springer Verlag, **96** (1990).
[7] J. Dombrowski; *Tridiagonal Matrix Representations of Cyclic Self-Adjoint Operators*, Pacific Jounral of Mathematics, **114**, No. 2 (1984), 325-334.
[8] J. Dombrowski; *Tridiagonal Matrix Representations of Cyclic Self-Adjoint Operators. II*, Pacific Jounral of Mathematics, **120**, No. 1 (1985), 47-53.
[9] R. Douglas; *Banach Algebra Techniques in Operator Theory, 2nd Edition*, Graduate Texts in Mathematics, Springer Verlag, **179**, 1998.
[10] P. L. Duren; *Extension of a Result of Beurling on Invariant Subspaces*, Transactions of the American Mathematical Society, **99** No. 2 (1961), 320-324.
[11] P. L. Duren; *Invariant Subspaces of Tridiagonal Operators*, Duke Math. J, **30** No. 2 (1963), 239-248.
[12] T. Furuta; *Invitation to Linear Operators*, Taylor & Francis, 2001.
[13] S. Haykin; *Neural Networks: A Comprehensive Foundation*, Macmillan College Publ. Co., 1994.
[14] K. Cox, P. Adams; *Formation of New Connections by a Generalisation of Hebbian Learning*, preprint, 2001.
[15] P. Hartman; *Ordinary Differential Equations*, Wiley, 1964.
[16] J. Hertz, A. Krogh, R. Palmer; *Introduction to the Theory of Neural Computation*, A Lecture Notes Volume, Santa Fe Institute Studies in the Sciences of Complexity, 1991.
[17] E. Oja; *A Simplified Neuron Model as a Principal Component Analyzer*, J. of Math. Biology, **15** (1982), 267-273.
[18] E. Oja, J. Karhunen; *A Stochastic Approximation of the Eigenvectors and Eigenvalues of the Expectation of a Random Matrix*, J. of Math. Analysis and Appl., **106** (1985), 69-84.
[19] J. Ringrose; *Compact Non-Self-Adjoint Operators*, Van Nostrand ReinHold Mathematical Studies **35**, 1994.

Fernanda Botelho
The University of Memphis, Mathematical Sciences Dept., Dunn Hall 373, 3721 Norriswood St., Memphis, TN 38152-3240, USA
E-mail address: mbotelho@memphis.edu

Annita Davis
The University of Memphis, Mathematical Sciences Dept., Dunn Hall 373, 3721 Norriswood St., Memphis, TN 38152-3240, USA
E-mail address: adavis2@memphis.edu