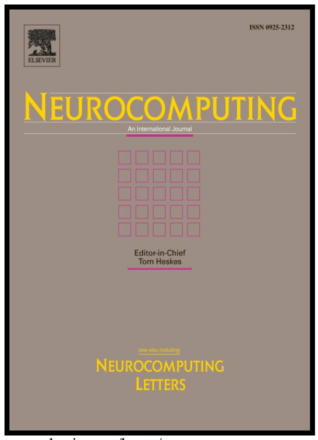
Author's Accepted Manuscript



www.elsevier.com/locate/neucom

A recursive least square algorithm for online kernel principal component extraction

João B.O. Souza Filho, Paulo S.R. Diniz

PII: S0925-2312(16)31507-7
DOI: http://dx.doi.org/10.1016/j.neucom.2016.12.031
Reference: NEUCOM17849

To appear in: *Neurocomputing*

Received date: 20 May 2016
Revised date: 23 November 2016
Accepted date: 8 December 2016

Cite this article as: João B.O. Souza Filho and Paulo S.R. Diniz, A recursive least square algorithm for online kernel principal component extraction *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2016.12.031

A recursive least square algorithm for online kernel principal component extraction

João B. O. Souza Filho[a,b], Paulo S. R. Diniz[a,c]

(e-mails: jbfilho@poli.ufrj.br, diniz@smt.ufrj.br)

[a]*Department of Electronics and Computer Engineering, Polytechnic School, Federal University of Rio de Janeiro, Avenida Athos da Silveira Ramos, 149, Technological Center, Building H, 2nd floor, rooms H-219(20) and 221, Rio de Janeiro, Brazil*
[b]*Electrical Engineering Postgraduate Program (PPEEL), Federal Center of Technological Education Celso Suckow da Fonseca, Avenida Maracanã, 229, Building E, 5th floor, Rio de Janeiro, Brazil*
[c]*Electrical Engineering Program (PEE), Alberto Luiz Coimbra Institute (COPPE), Federal University of Rio de Janeiro*

Abstract

The online extraction of kernel principal components has gained increased attention, and several algorithms proposed recently explore kernelized versions of the generalized Hebbian algorithm (GHA) [1], a well-known principal component analysis (PCA) extraction rule. Consequently, the convergence speed of such algorithms and the accuracy of the extracted components are highly dependent on a proper choice of the learning rate, a problem dependent factor. This paper proposes a new online fixed-point kernel principal component extraction algorithm, exploring the minimization of a recursive least-square error function, conjugated with an approximated deflation transform using component estimates obtained by the algorithm, implicitly applied upon data. The proposed technique automatically builds a concise dictionary to expand kernel components, involves simple recursive equations to dynamically define a specific learning rate to each component under extraction, and has a linear computational complexity regarding dictionary size. As compared to state-of-art kernel principal component extraction algorithms, results show improved convergence speed and accuracy of the components produced by the proposed method in five open-access databases.

Keywords:
kernel principal components analysis, kernel methods, online kernel algorithms, machine learning, generalized Hebbian algorithm

1. Introduction

Kernel principal components analysis (KPCA) [2] is a simple but powerful nonlinear generalization of the widely used PCA technique [3]. Originally stated as a Gram-matrix eigendecomposition problem [2], thus solvable by classical linear algebra methods [4], this technique faces problems with large scale datasets, for which the computational burden involved in Gram-matrix construction and factorization may turn the extraction process infeasible.

To address these problems, several authors have proposed incremental [5, 6, 7] and more recently online kernel component extraction algorithms [8, 9, 10]. Some examples are the online kernel Hebbian algorithm (OKHA) [8] and the subset kernel Hebbian algorithm (SubKHA) [9], which are extensions of the kernel Hebbian algorithm (KHA) [6]. In both cases, component estimates are expanded using concise and dynamically built dictionaries, but following different rules for mananging them. The kernel Hebbian algorithm (KHA) [6] is a nonlinear extension of the generalized Hebbian algorithm (GHA) [1].

A practical issue faced when using these algorithms is the choice of the learning rate factor, which critically affects the convergence speed and accuracy of extracted components. Usually, such relevant factor is set the same for all components under extraction by trial-and-error or following some meta-heuristic procedure. In case of KHA, the work [11] showed that adopting individual learning rates to each component under extraction results in better convergence and accuracy [11]. Nonetheless, the meta-heuristic procedure explored in this work is computationally expensive and requires the adjustment of several experimental parameters. In a similar way, Tanaka [10] has proposed a recursive least-square online kernel PCA extraction algorithm based on the iterative weighted rule (IWR) [12]. The strategy adopted by this rule is somewhat similar to the one proposed in [13] for the extraction of principal components, resulting in learning rates dynamically adjusted to each component under extraction. The accuracy of this algorithm, however, seems to be dependent on the choice of the weighting constants [13].

Motivated by these drawbacks this paper proposes a low-complexity online kernel principal component extraction algorithm, named recursive least square kernel Hebbian algorithm (RLS-KHA). The algorithm employs individual learning rates to each component under extraction, automatically tuned using a very simple iterative equation. The proposed RLS-KHA possesses iterative equations similar to KHA, and furthermore demands the setting of only one experimental parameter: the forgetting factor, exhibiting improved accuracy and convergence speed.

2. Kernel PCA

The KPCA [2] is a natural nonlinear extension of the PCA [3]. This technique implicitly produces mappings of real-