

控制与决策

Control and Decision

基于共轭梯度搜索的广义特征对追踪算法

蔡浩源, 陈捷, 张利军

引用本文:

蔡浩源, 陈捷, 张利军. 基于共轭梯度搜索的广义特征对追踪算法[J]. *控制与决策*, 2023, 38(7): 1927–1934.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1832>

您可能感兴趣的其他文章

Articles you may be interested in

[基于分布约束的非对称度量学习无监督行人重识别](#)

Asymmetric metric learning approach based on distribution constraints for unsupervised person re-identification

控制与决策. 2023, 38(6): 1703–1711 <https://doi.org/10.13195/j.kzyjc.2021.1598>

[基于改进秃鹰搜索算法的同步优化特征选择](#)

Simultaneous feature selection optimization based on improved bald eagle search algorithm

控制与决策. 2022, 37(2): 445–454 <https://doi.org/10.13195/j.kzyjc.2020.1025>

[基于可变形卷积的孪生网络目标跟踪算法](#)

Target tracking based on deformable convolution siamese network

控制与决策. 2022, 37(8): 2049–2055 <https://doi.org/10.13195/j.kzyjc.2021.0088>

[基于广义罚函数可行性准则的DE算法对不确定数据的处理](#)

Application of improved DE algorithm based on generalized penalty function feasibility criteria in uncertain data processing

控制与决策. 2021, 36(2): 498–504 <https://doi.org/10.13195/j.kzyjc.2019.0728>

[基于核典型相关性-熵成分分析的工业过程质量监测方法](#)

A quality monitoring method for industrial process based on kernel canonical correlation-entropy component analysis

控制与决策. 2021, 36(12): 2999–3006 <https://doi.org/10.13195/j.kzyjc.2020.0548>

基于共轭梯度搜索的广义特征对追踪算法

蔡浩源, 陈捷[†], 张利军

(西北工业大学航海学院, 西安 710072)

摘要: 研究广义特征对追踪算法, 通过探索基于共轭梯度搜索的标准特征向量追踪算法, 将其引入到广义特征对的提取. 所提算法具有自适应步长机制, 使不同特征搜索方向上的广义瑞利熵达到最优, 并适用于提取平稳矩阵束和非平稳矩阵束的广义特征对. 数值仿真中将所提算法与多个自适应广义特征向量提取算法进行了比较, 实验结果验证了所提算法的有效性.

关键词: 共轭梯度算法; 子空间追踪; 广义特征对; 特征提取; 广义特征向量; 广义瑞利熵

中图分类号: TP301 文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1832

引用格式: 蔡浩源, 陈捷, 张利军. 基于共轭梯度搜索的广义特征对追踪算法[J]. 控制与决策, 2023, 38(7): 1927-1934.

Generalized eigen-pairs tracking based on conjugate gradient method

CAI Hao-yuan, CHEN Jie[†], ZHANG Li-jun

(School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Generalized eigenvalue decomposition plays a vital role in statistical signal processing. Generalized decomposition aims to enhance the signal by seeking the directions that capture most of signal component but are orthogonal to the spaces constituting the noise component. Each generalized eigenvalue represents the optimal signal-to-noise ratio that can be obtained by projecting an observation into the corresponding eigen-direction. This paper proposes a generalized eigen-pairs tracking method based on conjugate gradient searching. The proposed method is variable step-size that seeks the generalized eigenvector in a sense that generalized Rayleigh quotient is optimal in the corresponding searching direction. It is suitable for extracting generalized eigenvectors from stationary and non-stationary matrix pencil. We compare the proposed method with multiple adaptive generalized extraction algorithms. The effectiveness of the proposed method is validated via numerical simulations.

Keywords: conjugate gradient method; subspace tracking; generalized eigen-pairs; feature extraction; generalized eigenvectors; generalized Rayleigh quotient

0 引言

提取广义特征向量 (generalized eigenvectors, GEV) 和特征值在众多信号处理和机器学习应用中有着举足轻重的作用, 例如阵列信号处理^[1]、语音信号增强^[2-3]、模式识别^[4-5] 以及盲源信号分离^[6].

广义特征分解的应用场景分为在线和离线. 离线场景的关键点在于利用合适的数值方法求解广义正交向量, 经典矩阵分解技术例如反幂迭代、Lanczos 和 Jacobi-Davidson 等批量处理方法^[7] 可实现该目的. 其中 Jacobi-Davidson 方法基于对矩阵对进行全分解, 计算复杂度较高. 基于 Krylov 子空间的 Lanczos 方法虽然避免了对矩阵对进行全分解, 但其存在数

值稳定性的问题, 且难以并行运算^[8]. 在统计信号处理等在线场景下, 关键点在于挖掘信息向量和先验观测之间的迭代关系, 以便降低计算复杂度和保证算法稳健性. 文献[9]利用了经典梯度下降法来迭代求取广义特征向量, 但通常面临着步长难以选择的难题. 文献[10]通过扩展经典的投影近似子空间追踪 (PAST)^[11] 的代价函数, 提出类递归最小二乘 (RLS) 的算法, 提高了算法的收敛速度. 在提高算法的稳健性方面, 一类鲁棒的拟牛顿 (quasi-Newton) 算法被提出, 这些算法通过耦合估计广义特征向量和广义特征值以实现特征对的提取^[12-13]. 神经网络学派提出的自稳定算法也获得了一定的关注^[14-16], 自稳定算

收稿日期: 2021-10-25; 录用日期: 2022-03-28.

基金项目: 国家自然科学基金项目 (62171380).

责任编辑: 孙宗耀.

[†]通讯作者. E-mail: jie.chen@nwpu.edu.cn.

法所追踪的权值向量的范数具有趋向单位长度的特性^[14,16],此外该类型算法形式简洁,具有节省处理器资源的优点.上述大部分算法仍依赖于常数步长的选取,而合适的步长往往依赖于对协方差统计量先验信息的观测.

不依赖步长的并行算法也是广义特征分解问题的研究热点之一.文献[17]提出在旧权值向量和实时观测向量的张成子空间中进行特征成分提取,但搜索向量维数小于信号子空间维数时,该算法收敛速度趋于缓慢.文献[18]提出将广义特征分解问题转换为标准的特征值问题,并用子空间迭代的方法进行求解,该算法具有较快的收敛速度,增强了其在实际工程的适用性.文献[19]通过随机采样技术^[20],提出了有效的高维矩阵的广义特征向量提取方法,适用于大规模流数据下的特征追踪.文献[20-21]提出了一类不动点迭代方程,这类方法兼具较快的计算速度和收敛速度.

共轭梯度(CG)算法在求解线性方程中被广泛研究.受到其自适应步长机制,良好的收敛特性和数值稳健等优点的启发,本文研究追踪标准特征向量的共轭梯度算法^[22],并将其引入到对广义特征对的追踪.

1 模型建立

在上述应用中,通常面临求解满足下列条件的特征分解表达式:

$$\begin{aligned} \mathbf{R}_y \mathbf{u}_i &= \lambda_i \mathbf{R}_v \mathbf{u}_i, \\ \mathbf{u}_i^H \mathbf{R}_v \mathbf{u}_j &= \delta_{ij}, \quad i, j \in \{1, \dots, N\}. \end{aligned} \quad (1)$$

其中: $\mathbf{R}_y, \mathbf{R}_v \in \mathbf{C}^{N \times N}$, \mathbf{u}_i 为第 i 个广义特征向量, $\lambda_1 \geq \dots \geq \lambda_N$ 为按大小排列的广义特征值, δ_{ij} 为克罗内克量.在信号处理中, \mathbf{R}_y 通常为观测信号的协方差矩阵, \mathbf{R}_v 为噪声及干扰信号的协方差矩阵. \mathbf{R}_y 和 \mathbf{R}_v 是自共轭的正定矩阵,因此其广义特征值非负.

考虑下述线性信号模型:

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{v}(k). \quad (2)$$

其中: $\mathbf{y}(k)$ 为 k 时刻的观测信号, $\mathbf{x}(k)$ 为 k 时刻的纯净信号, $\mathbf{v}(k)$ 为 k 时刻的加性噪声.通常情况下,假设纯净信号 $\mathbf{x}(k)$ 和加性噪声 $\mathbf{v}(k)$ 是不相关的,即 $E(\mathbf{x}\mathbf{v}^H) = \mathbf{O}_{N \times N}$.上述信号的协方差矩阵可分别定义为 $\mathbf{R}_y = E(\mathbf{y}\mathbf{y}^H)$, $\mathbf{R}_x = E(\mathbf{x}\mathbf{x}^H)$ 以及 $\mathbf{R}_v = E(\mathbf{v}\mathbf{v}^H)$.联合对角化是提取矩阵束 $(\mathbf{R}_y, \mathbf{R}_v)$ 广义特征向量的重要技术手段:

$$\begin{aligned} U^H \mathbf{R}_y U &= \Lambda, \\ U^H \mathbf{R}_v U &= I_N. \end{aligned} \quad (3)$$

上述等式等价于广义特征分解

$$\mathbf{R}_y U = \mathbf{R}_v U \Lambda. \quad (4)$$

假设信号子空间由广义特征矩阵 U 的 r 维主导向量张成,则观测数据在该 r 维主导子空间上投影后可保留纯净信号的成分.因此,利用 U 的 r 维主导子空间构建的滤波器,可使纯净信号“无失真”通过,并尽可能压制噪声能量,该滤波器在滤波结果上等价于最小方差无失真滤波器.若利用最大广义特征值对应的广义特征向量进行滤波器构建,则在牺牲一部分纯净信号能量的情况下最大限度压制噪声成分,该滤波器等价于最大信噪比滤波器.

在多数情况下,信号子空间和噪声子空间是时变的.为了达到追踪的目的,通常对协方差矩阵进行实时更新:

$$\begin{aligned} \hat{\mathbf{R}}_y(k+1) &= \\ &\beta \hat{\mathbf{R}}_y(k) + (1-\beta) \mathbf{y}(k+1) \mathbf{y}^H(k+1), \\ \hat{\mathbf{R}}_v(k+1) &= \\ &\alpha \hat{\mathbf{R}}_v(k) + (1-\alpha) \mathbf{v}(k+1) \mathbf{v}^H(k+1). \end{aligned} \quad (5)$$

其中 α 和 β 称为遗忘因子或平滑因子,取值在 0 和 1 之间.在追踪快速时变信号时, α 和 β 通常设定为较小值,以便更快地遗忘旧信息.而在追踪慢速时变信号时, α 和 β 通常取较大值,以减小观测数据的波动.

2 算法描述

本节描述基于共轭梯度的广义特征对追踪算法.为了叙述方便,首先对相关变量进行定义.

$\lambda_i(k)$: k 时刻第 i 个广义特征值预测值;

$r_i(k)$: k 时刻第 i 个广义特征向量的预测残差;

$\mathbf{u}_i(k)$: k 时刻第 i 个归一化的权值向量;

$\mathbf{x}_i(k)$: k 时刻第 i 个非归一化的权值向量;

$t_i(k)$: k 时刻第 i 个非归一化的权值向量更新步长;

$\mathbf{p}_i(k)$: k 时刻第 i 个广义特征向量的共轭搜索方向;

$q_i(k)$: k 时刻第 i 个广义特征向量的共轭更新步长.

2.1 基于共轭梯度算法的多广义特征对提取方法

首先考虑一个给定的矩阵束 $(\mathbf{R}_y, \mathbf{R}_v)$.假定 $\mathbf{u}_i(k)$ 为 k 时刻第 i 个归一化的权值向量,则 k 时刻第 i 个广义特征值可以预测为

$$\lambda_i(k) = J_i(k) = \mathbf{u}_i^H(k) \mathbf{R}_y \mathbf{u}_i(k). \quad (6)$$

根据式(1),可知 k 时刻第 i 个广义特征分解表达式的预测残差为

$$\mathbf{r}_i(k) = \lambda_i(k)\mathbf{R}_v\mathbf{u}_i(k) - \mathbf{R}_y\mathbf{u}_i(k). \quad (7)$$

若 $\mathbf{p}_i(k)$ 表示 \mathbf{u}_i 在第 k 次迭代的搜索方向, 则可根据如下方法更新广义特征向量的预测值:

$$\mathbf{x}_i(k+1) = \mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k), \quad (8)$$

其中 $t_i(k)$ 表示权值向量 $\mathbf{x}_i(k+1)$ 的更新步长. $t_i(k)$ 的选定可以最大化或最小化广义瑞利熵, 进而决定提取广义主成分或次成分. 简单地将 $t_i(k)$ 选定为常数值虽然可以满足提取平稳矩阵束广义特征对的需求, 但其大小的选择将难以平衡收敛速度和算法的稳健性. 通常, 当预测残差 $\mathbf{r}_i(k)$ 较大时, $t_i(k)$ 应自动选取为较大值; 当预测残差 $\mathbf{r}_i(k)$ 很小或 $\mathbf{u}_i(k)$ 接近收敛时, $t_i(k)$ 应自动选取为较小值, 以便降低对平衡状态的扰动, 这是一种“自适应”步长的思想, $t_i(k)$ 的选取将在下面进行详细推导. 注意到相邻时刻的共轭搜索方向是关于 \mathbf{R}_y 正交的, 即

$$\mathbf{p}_i^H(k+1)\mathbf{R}_y\mathbf{p}_i(k) = 0. \quad (9)$$

根据残差 $r_i(k)$ 的大小, 对新时刻的搜索方向进行更新

$$\mathbf{p}_i(k+1) = \mathbf{r}_i(k) - q_i(k)\mathbf{p}_i(k). \quad (10)$$

将上式代入到式(9)可得

$$\begin{aligned} \mathbf{p}_i^H(k+1)\mathbf{R}_y\mathbf{p}_i(k) &= \\ [\mathbf{r}_i(k) - q_i(k)\mathbf{p}_i(k)]^H\mathbf{R}_y\mathbf{p}_i(k) &= 0. \end{aligned} \quad (11)$$

式(11)进一步变成

$$q_i^*(k) = \frac{\mathbf{r}_i^H(k)\mathbf{R}_y\mathbf{p}_i(k)}{\mathbf{p}_i^H(k)\mathbf{R}_y\mathbf{p}_i(k)}, \quad (12)$$

其中 $(\cdot)^*$ 表示共轭符号. 因此, $q_i(k)$ 选择为

$$q_i(k) = \left[\frac{\mathbf{r}_i^H(k)\mathbf{R}_y\mathbf{p}_i(k)}{\mathbf{p}_i^H(k)\mathbf{R}_y\mathbf{p}_i(k)} \right]^*, \quad (13)$$

可保证相邻时刻搜索方向是自共轭的. 为了提取多个广义特征对, 可以采用 Gram-Schmidt 正交化方法

$$\mathbf{u}_i(k+1) = \frac{\left[\mathbf{I}_N - \sum_{j=1}^{i-1} \mathbf{u}_j(k+1)\mathbf{u}_j^H(k+1) \right]^H \mathbf{R}_v \mathbf{x}_i(k+1)}{\sqrt{[\mathbf{x}_i(k+1)]^H \left\{ \mathbf{R}_v - \sum_{j=1}^{i-1} \mathbf{R}_v \mathbf{u}_j(k+1)\mathbf{u}_j^H(k+1) \right\} \mathbf{x}_i(k+1)}}. \quad (20)$$

定义

$$\mathbf{B}_i = \left\{ \mathbf{I}_N - \sum_{j=1}^{i-1} \mathbf{u}_j(k+1)\mathbf{u}_j^H(k+1) \right\}, \quad (21)$$

$$\mathbf{C}_i = \left\{ \mathbf{R}_v - \sum_{j=1}^{i-1} \mathbf{R}_v \mathbf{u}_j(k+1)\mathbf{u}_j^H(k+1) \right\}, \quad (22)$$

进行正交投影. 需要注意的是, 期望提取的权值向量 $\mathbf{u}_i(k+1)$ 并不满足 $\mathbf{u}_i^H(k+1)\mathbf{R}_v\mathbf{u}_i(k+1) = 1$ 的约束, 因此需要在标准正交化的基础上进行适当改进. 假定归一化权值向量 $\mathbf{u}_1(k), \dots, \mathbf{u}_N(k)$ 满足如下约束:

$$[\mathbf{u}_i(k)]^H \mathbf{R}_v \mathbf{u}_j(k) = \delta_{ij}, \quad i, j \in \{1, \dots, N\}. \quad (14)$$

$\mathbf{x}_i(k+1)$ 通过归一化可得到

$$\mathbf{u}_i(k+1) = \frac{\mathbf{x}_i(k+1)}{\sqrt{[\mathbf{x}_i(k+1)]^H \mathbf{R}_v \mathbf{x}_i(k+1)}}. \quad (15)$$

因此, 第 $i+1$ 个非归一化的权值向量可构造为

$$\begin{aligned} \mathbf{u}'_i(k+1) &= \\ \mathbf{x}_i(k+1) - \frac{\langle \mathbf{u}_1(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v}}{\langle \mathbf{u}_1(k+1), \mathbf{u}_1(k+1) \rangle_{\mathbf{R}_v}} \mathbf{u}_1(k+1) - \\ &\quad \frac{\langle \mathbf{u}_2(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v}}{\langle \mathbf{u}_2(k+1), \mathbf{u}_2(k+1) \rangle_{\mathbf{R}_v}} \mathbf{u}_2(k+1) - \dots - \\ &\quad \frac{\langle \mathbf{u}_{i-1}(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v}}{\langle \mathbf{u}_{i-1}(k+1), \mathbf{u}_{i-1}(k+1) \rangle_{\mathbf{R}_v}} \mathbf{u}_{i-1}(k+1). \end{aligned} \quad (16)$$

可简化为

$$\begin{aligned} \mathbf{u}'_i(k+1) &= \\ \mathbf{x}_i(k+1) - \langle \mathbf{u}_1(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v} \mathbf{u}_1(k+1) - \\ &\quad \langle \mathbf{u}_2(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v} \mathbf{u}_2(k+1) - \dots - \\ &\quad \langle \mathbf{u}_{i-1}(k+1), \mathbf{x}_i(k+1) \rangle_{\mathbf{R}_v} \mathbf{u}_{i-1}(k+1). \end{aligned} \quad (17)$$

进一步表示为

$$\begin{aligned} \mathbf{u}'_i(k+1) &= \\ \left[\mathbf{I}_N - \sum_{j=1}^{i-1} \mathbf{u}_j(k+1)\mathbf{u}_j^H(k+1) \right]^H \mathbf{R}_v \mathbf{x}_i(k+1). \end{aligned} \quad (18)$$

因此, 归一化的权值向量为

$$\mathbf{u}_i(k+1) = \frac{\mathbf{u}'_i(k+1)}{\sqrt{[\mathbf{u}'_i(k+1)]^H \mathbf{R}_v \mathbf{u}'_i(k+1)}}. \quad (19)$$

通过进一步推导可得到

其中 \mathbf{C}_i 被称为正交补投影矩阵, 其作用是使已提取的特征空间正交于当前搜索空间, 从而实现非主导广义特征向量的逐个提取(证明参见文献[22]).

根据上述定义, $\mathbf{u}_i(k+1)$ 可进一步表示为

$$\mathbf{u}_i(k+1) = \frac{\mathbf{B}_i \mathbf{x}_i(k+1)}{\sqrt{[\mathbf{x}_i(k+1)]^H \mathbf{C}_i \mathbf{x}_i(k+1)}}. \quad (23)$$

2.2 权值向量的步长更新方法

本节以优化广义瑞利熵为目标,推导权值向量步长 $t_i(k)$ 的更新公式. 根据式(8)可得

$$\begin{aligned} J_i(k+1) &= \\ \mathbf{u}_i^H(k+1)\mathbf{R}_y\mathbf{u}_i(k+1) &= \\ \frac{\mathbf{x}_i^H(k+1)\mathbf{B}_i^H\mathbf{R}_y\mathbf{B}_i\mathbf{x}_i(k+1)}{\mathbf{x}_i^H(k+1)\mathbf{C}_i\mathbf{x}_i(k+1)} &= \\ \frac{[\mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k)]^H\mathbf{B}_i^H\mathbf{R}_y\mathbf{B}_i[\mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k)]}{[\mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k)]^H\mathbf{C}_i[\mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k)]} &. \end{aligned} \quad (24)$$

定义以下变量以便进一步化简:

$$\mathbf{u}_b = \mathbf{B}_i\mathbf{u}_i(k), \quad (25)$$

$$\mathbf{u}_p = \mathbf{B}_i\mathbf{p}_i(k), \quad (26)$$

$$\lambda = \mathbf{u}_b^H\mathbf{R}_y\mathbf{u}_b, \quad (27)$$

$$a = \mathbf{u}_b^H\mathbf{R}_y\mathbf{u}_p, \quad (28)$$

$$b = \mathbf{u}_p^H\mathbf{R}_y\mathbf{u}_p, \quad (29)$$

$$e = \mathbf{u}_i^H(k)\mathbf{C}_i\mathbf{u}_i(k), \quad (30)$$

$$c = \mathbf{u}_i^H(k)\mathbf{C}_i\mathbf{p}_i(k), \quad (31)$$

$$d = \mathbf{p}_i^H(k)\mathbf{C}_i\mathbf{p}_i(k). \quad (32)$$

因此,广义瑞利熵可写成如下形式:

$$J_i(k+1) = \frac{\lambda + ta + t^*a^* + tt^*b}{e + tc + t^*c^* + tt^*d}. \quad (33)$$

若 $t_i(k)$ 的选取使得广义瑞利熵达到最优(即瑞利熵最大化或最小化),则有

$$\frac{\partial J_j(k+1)}{\partial t_i(k)} = 0. \quad (34)$$

容易验证,当上式成立时有

$$\begin{aligned} (bc - ad)t_i^2(k) + (be - \lambda d + a^*c - \\ ac^*)t_i(k) + (a^*e - \lambda c^*) = 0. \end{aligned} \quad (35)$$

上式为二次函数,所以 $t_i(k)$ 的解为

$$t_i(k) = \frac{-f \pm \sqrt{f^2 - 4hg}}{2h}. \quad (36)$$

其中

$$f = be - \lambda d + a^*c - ac^*, \quad (37)$$

$$g = a^*e - \lambda c^*, \quad (38)$$

$$h = bc - ad. \quad (39)$$

经过进一步计算将 $J_i(k+1)$ 化成

$$\begin{aligned} J_i(k+1) &= \\ J_i(k+1)(1-e) &+ \end{aligned}$$

$$\begin{aligned} \frac{t(ae - \lambda c) + t^*(a^*e - \lambda c^*) + tt^*(be - \lambda d)}{\mathbf{x}_i^H(k+1)\mathbf{C}_i\mathbf{x}_i(k+1)} + \lambda &\triangleq \\ J_i(k+1)(1-e) + \Gamma + \lambda. \end{aligned} \quad (40)$$

由于 $\lambda > 0$,假定要使 $J_i(k+1)$ 最大化,则第2项 Γ 的值应该为正数;欲使 $J_i(k+1)$ 最小化, Γ 的值应该为负数. 因此可断定, $t_i(k)$ 中根号项前符号的选取与 Γ 的符号有关. 进一步对 Γ 单独研究,由式(36)可得

$$a^*e - \lambda c^* = -t^2(bc - ad) - t(be - \lambda d + a^*c - ac^*). \quad (41)$$

将(41)代入到式(40)中 Γ 的表达式,可得

$$\begin{aligned} \Gamma &= - \frac{tt^*}{2\mathbf{x}_i^H(k+1)\mathbf{C}_i\mathbf{x}_i(k+1)} \times \\ &\{2t^*(bc^* - a^*d) + (be - \lambda d + ac^* - a^*c) + \\ &2t(bc - ad) + (be - \lambda d + a^*c - ac^*)\}. \end{aligned} \quad (42)$$

上式括号内有两个共轭项,其结果是实数. 其中一项有

$$\begin{aligned} 2t(bc - ad) + (be - \lambda d + a^*c - ac^*) &= \\ \pm \sqrt{(be - \lambda d + a^*c - ac^*) - 4(bc - ad)(a^*e - \lambda c^*)} &\triangleq \\ \pm (\zeta + j\xi), \end{aligned} \quad (43)$$

其中 ζ 和 ξ 分别表示实部和虚部. 同时令

$$\alpha = \frac{tt^*}{\mathbf{x}_i^H(k+1)\mathbf{C}_i\mathbf{x}_i(k+1)}. \quad (44)$$

所以 $J_i(k+1)$ 可最终表示成

$$J_i(k+1) = J_i(k+1)(1-e) - \alpha(\pm\zeta) + \lambda. \quad (45)$$

由于 $\alpha > 0$,最大化广义瑞利熵应该选择 ξ 的相反符号,而最小化广义瑞利熵应该选择 ξ 的相同符号. 由此可知,步长

$$t_i(k) = \frac{-f - \sqrt{f^2 - 4hg}}{2h}$$

是使瑞利熵最大化的最优选择,而步长

$$t_i(k) = \frac{-f + \sqrt{f^2 - 4hg}}{2h}$$

是使瑞利熵最小化的最优选择.

假定真值广义特征向量为 $\mathbf{v}_1, \dots, \mathbf{v}_N$, $t_i(k)$ 的选取实质上保证了 $\mathbf{x}_i(k+1) = \mathbf{u}_i(k) + t_i(k)\mathbf{p}_i(k)$ 的广义瑞利熵 $\mathbf{J}(\mathbf{x}_i(k+1)) > \mathbf{J}(\mathbf{u}_i(k)) = \mathbf{J}(\mathbf{x}_i(k))$, $t_i(k) = 0$ 对应于向量 $\mathbf{u}_i(k)$ 的广义瑞利熵. 当 $i = 1$ 时,为了保证权值向量 $\mathbf{u}_1(k)$ 只收敛到主导广义特征向量 \mathbf{v}_1 ,初始值 $\mathbf{u}_1(0)$ 一定要在 \mathbf{v}_1 对应的特征向量上有投影,即 $\langle \mathbf{u}_1(0), \mathbf{v}_1 \rangle_{\mathbf{R}_x} \neq 0$. 此时根据广义瑞利熵在 \mathbf{v}_1 附近的单调性, k 时刻的 $\mathbf{u}_1(k)$ 在 \mathbf{v}_1 对应的特征向量上有投影分量一定不为零,否则将出现 $\mathbf{J}(\mathbf{x}_i(k+1)) < \mathbf{J}(\mathbf{u}_1(k))$. 此时只要当 $\mathbf{u}(k)$ 在鞍点

av_2, \dots, av_{N-1} 上的投影分量还存在, 步长 $t_1(k)$ 总是使得 $\mathbf{J}(\mathbf{u}_1(k+1)) = \mathbf{J}(\mathbf{x}_1(k+1)) > \mathbf{J}(\mathbf{u}_1(k)) = \mathbf{J}(\mathbf{x}_1(k))$, 并且只有当 $\mathbf{u}_1(k)$ 在鞍点 av_2, \dots, av_{N-1} 上的投影分量收敛到 0 才有 $t_1(k) = 0$. $t_1(k)$ 的存在保证了 $\mathbf{x}_1(k+1)$ 是一个收敛到 \mathbf{v}_1 主导广义特征方向的序列, 即 $\mathbf{J}(\mathbf{x}_1(k+1)) > \mathbf{J}(\mathbf{x}_1(k)) > \dots > \mathbf{J}(\mathbf{x}_1(0))$. 简言之, 若 $\mathbf{u}_1(k)$ 在主导广义特征向量 \mathbf{v}_1 上存在投影成分, 则权向量 $\mathbf{x}_1(k+1)$ 不可能收敛至鞍点 av_2, \dots, av_{N-1} , 否则广义瑞利熵的代价值将持续递增直至跳出鞍点. 同理, 由于正交补矩阵(21)和(22)的存在, $\mathbf{u}_i(k+1)$ 和 $\mathbf{x}_i(k+1)$ 同样将收敛至 \mathbf{v}_i 对应的特征方向上.

2.3 算法总结

算法的具体实现过程总结如下

step 1: 输入协方差矩阵 \mathbf{R}_y 、 \mathbf{R}_v , 设定提取广义特征对数目 r , 初始化权向量 $\mathbf{u}_i(0)$, 初始化共轭搜索方向 $\mathbf{p}_i(0)$, 设置索引 $i = 1$, 设置迭代次数 $k = 1$, 迭代总数 N .

step 2: 若 $k > N$, 则跳出循环, 否则进行下列步骤.

step 3: 根据式(6)计算广义特征值 $\lambda_i(k)$.

step 4: 根据式(7)计算 $\mathbf{r}_i(k)$.

step 5: 根据 2.2 节所述策略计算步长 $t_i(k)$.

step 6: 根据式(8)更新 $\mathbf{x}_i(k+1)$.

step 7: 根据式(13)计算 $q_i(k)$.

step 8: 根据式(10)更新 $\mathbf{p}_i(k+1)$.

step 9: 根据式(20)计算归一化向量 $\mathbf{u}_i(k+1)$.

step 10: $i = i + 1$, 当 $i \leq r$ 返回 step 3; 否则 $i = 1$, $k = k + 1$, 返回 step 2.

下面分析所提算法提取主导广义特征对时的计算复杂度. 按照文献[13], 对主导运算量, 即矩阵乘矩阵的操作数进行分析. 计算第 i 个广义特征值 λ_i 需要 N^2 个乘法运算, 由于 $\mathbf{R}_y \mathbf{u}_i(k)$ 已通过式(6)计算, 计算第 i 个残差向量只需要 N^2 个乘法运算. 计算变步长 $t_i(k)$ 主要花费在等式(25)~(32)的计算上, 共需花费 $6N^2$ 个乘法运算. 归一化 $\mathbf{x}_i(k+1)$ 花费 N^2 个乘法运算. 因此, 所提算法提取主导广义特征向量的计算复杂度为 $9N^2 + \mathcal{O}(N)$.

3 数值仿真

本节对推导的算法进行数值仿真, 以验证算法的有效性. 在第 1 个实验中, 采用正弦叠加信号模型生成观测数据; 第 2 个实验考虑蜂窝无线网络的信号增

强问题, 在该场景下, 为了尽可能压制同波段干扰和噪声能量, 需要设计最大输出信噪比滤波器, 可以证明最大输出信噪比滤波器的构建依赖于主导广义特征向量的提取. 将所提算法与下述几种不同的算法进行比较:

1) quasi-Newton^[13]: 基于拟牛顿算法的广义特征对提取算法;

2) Power-like^[18]: 一种基于子空间迭代的广义特征向量提取算法;

3) R-GEVE^[17]: 基于降秩技术的广义特征向量提取算法;

4) Gradient^[9]: 基于梯度下降的广义特征向量提取算法.

此外, 为了衡量广义特征对的追踪精度, 引入对数坐标下的余弦相似度残差

$$DC_i(k) = 10 \log_{10} \left(1 - \frac{|\mathbf{u}_i^H(k) \mathbf{u}_i|}{\|\mathbf{u}_i(k)\| \|\mathbf{u}_i\|} \right). \quad (46)$$

所得结果经过 100 次蒙特卡洛仿真取平均.

3.1 正弦叠加信号模型

在该实验中, 观测向量的信号模型给定如下:

$$\begin{aligned} v(n) &= \sqrt{2} \sin(0.62\pi n + \theta_1) + \zeta_1(n), \\ y(n) &= \sqrt{2} \sin(0.46\pi n + \theta_2) + \\ &\quad \sqrt{2} \sin(0.74\pi n + \theta_3) + \zeta_2(n). \end{aligned} \quad (47)$$

其中: $\theta_i (i = 1, 2, 3)$ 为初始相位, 服从 $[0, 2\pi]$ 的均匀分布; $\zeta_1(n)$ 和 $\zeta_2(n)$ 为零均值的高斯白噪声, 其方差为 $\sigma_1^2 = \sigma_2^2 = 0.1$.

观测向量的维度为 $N = 8$, 即 $\mathbf{y}(k) = [y(k), \dots, y(k-N+1)]^T$ 以及 $\mathbf{v}(k) = [x(k), \dots, v(k-N+1)]^T$, 矩阵束的真值定义如下:

$$\begin{aligned} [\mathbf{R}_v]_{mn} &= \cos[0.62\pi(m-n)] + \delta_{mn} \sigma_1^2, \\ [\mathbf{R}_y]_{mn} &= \cos[0.46\pi(m-n)] + \\ &\quad \cos[0.74\pi(m-n)] + \delta_{mn} \sigma_2^2. \end{aligned} \quad (48)$$

在该实验中, 提取观测向量前 3 个主导广义特征对. 实验设置采用与文献[13]完全一致的参数: 协方差矩阵初始化为 $\hat{\mathbf{R}}_y(0) = \hat{\mathbf{R}}_v(0) = \mathbf{I}_N$, 设定遗忘因子 $\alpha = \beta = 0.998$. 对于依赖步长的算法, quasi-Newton 算法的步长设置为 $\alpha = \beta = 0.001$; Gradient 算法的步长设置为 $\alpha = \beta = -0.0001$. 权向量 $\mathbf{u}_1(0) = \mathbf{e}_1$, 其中 \mathbf{e}_1 为单位矩阵的第 i 个列向量. 此外, 为了保证 $\mathbf{u}_2(0)$ 和 $\mathbf{u}_3(0)$ 在相应的特征方向上有投影分量, 对它们进行随机初始化.

最后的实验结果展示如图1~图4. 其中: 图1~图3给出了5种算法在追踪前3个广义特征向量时的对数余弦相似度残差, 图4给出了能够提取广义特征值的两种算法在追踪前3个广义特征值的收敛曲线, 虚线表示真实广义特征值.

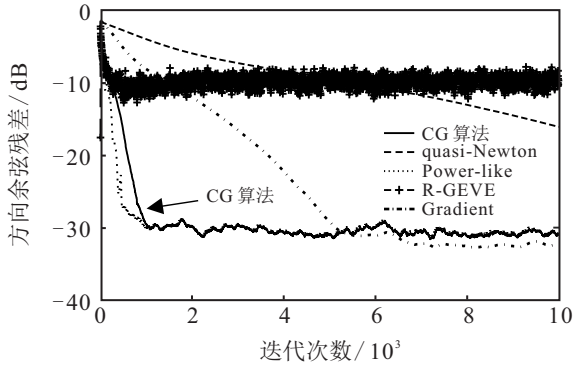


图1 第1个广义特征向量的收敛情况

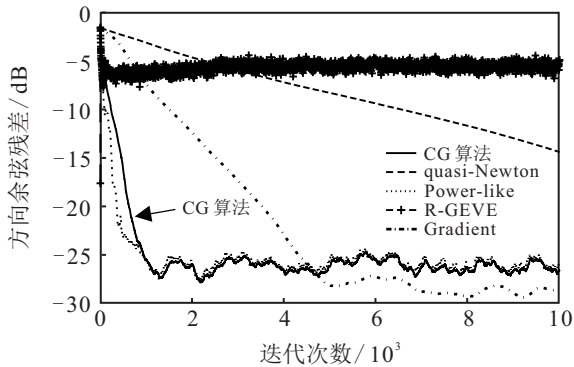


图2 第2个广义特征向量的收敛情况

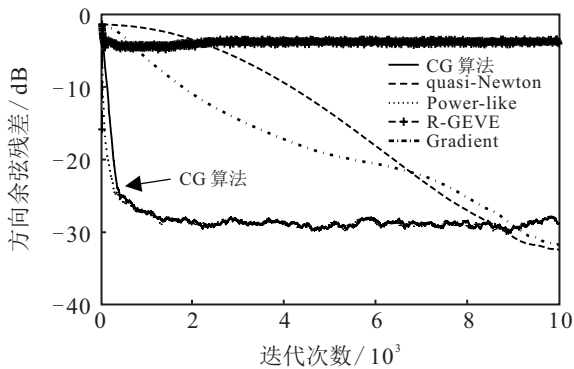


图3 第3个广义特征向量的收敛情况

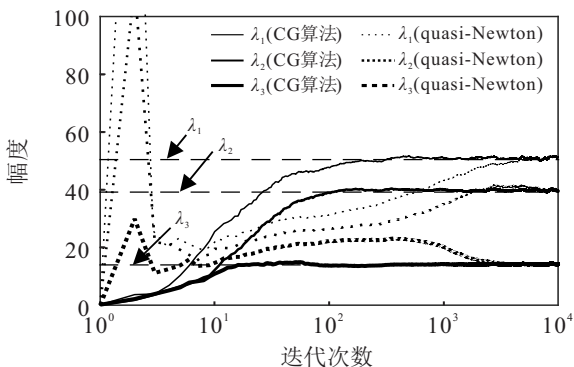


图4 广义特征值的收敛情况

从图1~图3可知: CG算法和Power-like算法的收敛速度最快, 但CG算法在初始收敛阶段比Power-like算法更稳健; R-GEVE算法虽然在提取主导广义特征向量时收敛到真值附近, 但在提取多个广义特征向量时明显面临着数值稳定性的问题, 导致算法最后不收敛; 基于固定步长的算法quasi-Newton和Gradient虽然在提取某个向量时稳态误差更低, 但其代价是付出大量的迭代次数. 在自适应信号处理的实际应用中, 通常难以选取合适的步长以取得预测精度和收敛速度之间的折衷, 严重限制了步长算法的实用性.

从图4可以看出, CG算法在追踪广义特征值时比固定步长下quasi-Newton收敛更快, 而后者收敛所需的迭代次数为 $5e^3$ 以上.

3.2 蜂窝无线网络中的信号增强应用

本节关注蜂窝网络应用中的信号增强问题. 由于同波段信号以及噪声的干扰, 严重降低了用户的通信质量. 为了使天线阵列形成一定的波束指向, 需要设计阵列的合适权向量. 最优的复权向量由主导广义特征向量构建而成, 而最优权值形成的自适应波束可最大限度地输出有用信号并压制干扰和噪声信号, 从而提高蜂窝网络的通信容量^[1]. 在此实验中, 输入信号 $\mathbf{y}(k)$ 和 $\mathbf{v}(k)$ 是定义在空域的复信号, 代表了发射信号的多径衰减路径, 即 $\mathbf{y}(k) = [y_1(k), \dots, y_N(k)]^T$ 和 $\mathbf{v}(k) = [v_1(k), \dots, v_N(k)]^T, k \geq N$, 其中 $y_i(k)$ 和 $v_i(k) (i = 1, \dots, N)$ 代表了入射到阵列第 i 个阵元的信号. 这里考虑 N 元的天线阵列接收到一个多径信号($q = 1$)以及一个多径干扰($q = 2$). 在信号模型中, 矩阵束 $(\mathbf{R}_y, \mathbf{R}_v)$ 的 (m, n) ($m, n = 1, \dots, N$)元素为

$$[\mathbf{R}_y]_{mn} = \frac{1}{I} \sum_{i=0}^{I-1} e^{j(\phi_{m,i}^{(1)} - \phi_{n,i}^{(1)})},$$

$$[\mathbf{R}_v]_{mn} = \delta_{mn} \sigma^2 + \frac{1}{I} \sum_{i=0}^{I-1} e^{j(\phi_{m,i}^{(2)} - \phi_{n,i}^{(2)})}. \quad (49)$$

其中

$$\phi_{m,i}^{(q)} = m\pi \sin[\theta_q - \Delta_q/2 + i\Delta_q/(I-1)],$$

$$q = 1, 2,$$

代表给定入射角度 θ_q 时的多径信号的相位; I 代表每个信号的路径总数; j 代表虚数单位; σ^2 代表复高斯白噪声的方差.

信号以及干扰均匀散布在夹角为 Δ_q 的每个扇区. 考虑 $N = 4$ 元的天线阵列, $I = 12, \Delta_1 =$

$30, \theta_2 = 15, \Delta_2 = 30, \sigma^2 = 0.01$ (对应信噪比 $SNR = 20 \text{ dB}$)^[12].

本实验中采用的衡量准则以及实验设置与3.1节保持一致. 图5展示了不同算法在提取主导广义特征向量时的收敛情况. 由图5可以看出Power-like和CG算法的收敛速度是最快的, 并且是quasi-Newton和Gradient算法的4到5倍. 步长算法花费接近 $1e^4$ 次迭代才接近收敛, 严重限制了其实际可用性. 此外, R-GEVE算法在该信号模型下同样面临着不收敛的难题.

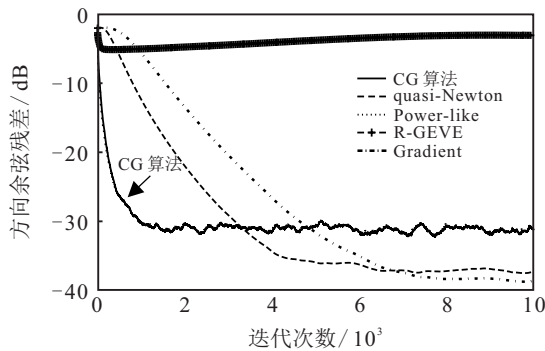


图5 主导广义特征向量的收敛情况

不同算法在提取主导广义特征向量(对)时的计算复杂度比较总结在表1, 虽然R-GEVE算法计算复杂度是最低的, 但CG算法更好地平衡了收敛速度和计算复杂度. 此外, 除了所提算法和quasi-Newton算法之外, Power-like、R-GEVE、Gradient的原始算法中均没有专门针对广义特征值设计提取策略.

表1 不同算法的计算复杂度比

算法	计算复杂度
CG 算法	$9N^2 + \mathcal{O}(N)$
quasi-Newton	$7N^2 + \mathcal{O}(N)$
Power-like	$13N^2 + \mathcal{O}(N)$
R-GEVE	$6N^2 + \mathcal{O}(N)$
Gradient	$10N^2 + \mathcal{O}(N)$

4 总结

本文推导了基于共轭梯度算法的广义特征对追踪方法, 该方法适用于追踪非平稳目标和平稳目标, 具有较低的计算复杂度和自适应步长机制. 本文在正弦叠加信号和蜂窝网络的信号模型中与不同算法进行了比较, 实验结果显示所提算法具有非常快的收敛速度, 很好地平衡了计算复杂度、步长选择、数值稳定性和收敛速度, 具有较高的工程实用性.

参考文献(References)

[1] Morgan D R. Downlink adaptive array algorithms for cellular mobile communications[J]. IEEE Transactions

on Communications, 2003, 51(3): 476-488.

[2] Zhao Y K, Chen J, Chen J D. Recursive variable span linear filter for noise reduction[J]. IEEE Signal Processing Letters, 2019, 26(12): 1902-1906.

[3] Benesty J, Christensen M G, Jensen J R. Signal enhancement with variable span linear filters[M]. Singapore: Springer Singapore, 2016.

[4] Hiraoka K, Hamahira M, Hidai K, et al. Fast algorithm for online linear discriminant analysis[J]. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2001, 84(6): 1431-1441.

[5] Mangasarian O L, Wild E W. Multisurface proximal support vector machine classification via generalized eigenvalues[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(1): 69-74.

[6] Tomé A M. The generalized eigendecomposition approach to the blind source separation problem[J]. Digital Signal Processing, 2006, 16(3): 288-302.

[7] Bai Z, Demmel J, Dongarra J, et al. Templates for the solution of algebraic eigenvalue problems: A practical guide[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2000: 1-5.

[8] Kaloorazi M F, de Lamare R C. Subspace-orbit randomized decomposition for low-rank matrix approximations[J]. IEEE Transactions on Signal Processing, 2018, 66(16): 4409-4424.

[9] Nguyen T D, Takahashi N, Yamada I. An adaptive extraction of generalized eigensubspace by using exact nested orthogonal complement structure[J]. Multidimensional Systems and Signal Processing, 2013, 24(3): 457-483.

[10] Yang J, Xi H S, Yang F, et al. RLS-based adaptive algorithms for generalized eigen-decomposition[J]. IEEE Transactions on Signal Processing, 2006, 54(4): 1177-1188.

[11] Yang B. Projection approximation subspace tracking[J]. IEEE Transactions on Signal Processing, 1995, 43(1): 95-107.

[12] Nguyen T D, Yamada I. Adaptive normalized quasi-Newton algorithms for extraction of generalized eigen-pairs and their convergence analysis[J]. IEEE Transactions on Signal Processing, 2013, 61(6): 1404-1418.

[13] Feng X W, Kong X Y, Duan Z S, et al. Adaptive generalized eigen-pairs extraction algorithms and their convergence analysis[J]. IEEE Transactions on Signal Processing, 2016, 64(11): 2976-2989.

[14] 汪奔, 孔祥玉, 冯晓伟, 等. 一种耦合的主/次特征对提取神经网络算法[J]. 信息与控制, 2017, 46(5):

536-542.
 (Wang B, Kong X Y, Feng X W, et al. A coupled principal/minor eigen-pairs extraction neural network algorithm[J]. Information and Control, 2017, 46(5): 536-542.)

[15] 冯晓伟, 孔祥玉, 马红光, 等. 自稳定的双目的特征对提取算法及其收敛性分析[J]. 控制与决策, 2017, 32(4): 600-606.
 (Feng X W, Kong X Y, Ma H G, et al. Unified self-stabilizing eigen-pairs extraction algorithm and its convergence analysis[J]. Control and Decision, 2017, 32(4): 600-606.)

[16] Feng X W, Kong X Y, Ma H G, et al. A novel unified and self-stabilizing algorithm for generalized eigenpairs extraction[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(12): 3032-3044.

[17] Attallah S, Abed-Meraim K. A fast adaptive algorithm for the generalized symmetric eigenvalue problem[J]. IEEE Signal Processing Letters, 2008, 15: 797-800.

[18] Tanaka T. Fast generalized eigenvector tracking based on the power method[J]. IEEE Signal Processing Letters, 2009, 16(11): 969-972.

[19] Cai H, Kaloorazi M F, Chen J, et al. Online dominant

generalized eigenvectors extraction via a randomized method[C]. The 28th European Signal Processing Conference (EUSIPCO). Amsterdam, 2021: 2353-2357.

[20] Cai H Y, Kaloorazi M F, Chen J. Online generalized eigenvectors extraction via a fixed-point approach[J]. IEEE Transactions on Signal Processing, 2021, 69: 2435-2451.

[21] Cai H Y, Zhao Y K, Chen J, et al. Online construction of variable span linear filters using a fixed-point approach[J]. IEEE Signal Processing Letters, 2021, 28: 404-408.

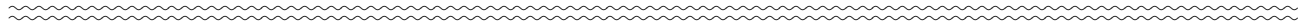
[22] Fu Z Q, Dowling E M. Conjugate gradient eigenstructure tracking for adaptive spectral estimation[J]. IEEE Transactions on Signal Processing, 1995, 43(5): 1151-1160.

作者简介

蔡浩源 (1995—), 男, 硕士生, 从事自适应信号处理的研究, E-mail: haoyuan.cai@nwpu.edu.cn;

陈捷 (1984—), 男, 教授, 博士生导师, 从事信号处理、机器学习等研究, E-mail: jie.chen@nwpu.edu.cn;

张利军 (1973—), 男, 教授, 博士生导师, 从事智能控制理论与应用、机器学习等研究, E-mail: zhanglj7385@nwpu.edu.cn.



下期要目

端边云协同的复杂工业过程运行控制智能系统 柴天佑, 等

复杂工业过程特征建模方法及应用研究 乔俊飞, 等

航天器姿轨控制研究综述: 微分几何控制方法 姜斌, 等

网络攻击下的信息物理系统安全状态估计研究综述 杨光红, 等

基于熵理论的综合能源系统不可能三角研究综述 孙秋野, 等

类严格反馈系统的模糊弱扰动解耦控制 刘晓平, 等

从结构推断到根因识别——工业过程故障根因诊断研究综述 赵春晖, 等

产品溯源研究综述及前景展望 黄敏, 等

基于多智能体的数字孪生及其在工业中应用的综述 张颖伟, 等

基于模糊强化学习的空气压缩机群组协调预测控制 王伟, 等

面向厚板板形质量监控与异常诊断的大数据可视分析 丁进良, 等

正则/非正则最优控制方法在最优趋同控制的应用 张焕水, 等

竞争环境下两个B2C服务共享平台的均衡决策 樊治平, 等

具有未知参数的非线性系统动态优化 付俊, 等