1. **Subspace Learning and Imputation for Streaming Big Data Matrices and Tensors**

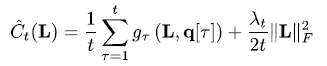
Author: Morteza Mardani, Gonzalo Mateos, Georgios B. Giannakis

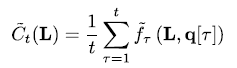
Other: IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2015

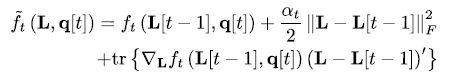
主要内容：

要解决的问题： C:\Users\690\AppData\Local\Temp\1554887291(1).png

目标函数：

C:\Users\690\AppData\Local\Temp\1554887887(1).png







证明思路：在三条假设下，首先证 是拟鞅序列，故收敛，且收敛到 C。之后证▽C也收敛，且收敛到0矩阵。

1. **Dynamic Anomalography: Tracking Network Anomalies Via Sparsity and Low Rank**

Author: Morteza Mardani, Gonzalo Mateos,Member, Georgios B. Giannakis

Other: IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, 2013

主要内容：在五条假设下（等价于上一篇中的三条假设），有Lemma1、2、3，证明之后可以推得分解之后的矩阵可收敛到平稳点（非局部最优，但几乎等价于局部最优）。

1. **Online Learning for Matrix Factorization and Sparse Coding**

Author: Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro

Other: INRIA - WILLOW Project-Team, University of Minnesota

Journal of Machine Learning Research, 2010

主要内容：与第一篇的证明思路大体相似，但证明过程稍有出入

以上三篇文章都是针对数据不断增加的过程进行研究的，数据矩阵和分解得到的左矩阵规模一直在变化

1. **Guaranteed Matrix Completion via Non-Convex Factorization**

Author: Ruoyu Sun and Zhi-Quan Luo

Other: Germany, IBM(USA)

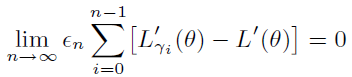
IEEE TRANSACTIONS ON INFORMATION THEORY, 2016

主要内容：文中提出可以从几何的角度证明MF的分解是唯一的,本质上还是使用局部凸性

1. **Large-Scale Matrix Factorization with Distributed Stochastic Gradient Descent**

Author: Rainer Gemulla, Peter J. Haas, Erik Nijkamp, Yannis Sismanis,

Other: KDD’11, 2011

主要内容：要求迭代中的步长是收敛的

在步长满足 的情况下,有

1. **Projected Gradient Methods for Nonnegative Matrix Factorization**

本质上还是递减步长

矩阵分解能到stationary point， 但是是最小二乘法

1. **Global Convergence of Stochastic Gradient Descent for Some non-convex Matrix Problems**

是矩阵分解问题，但不是被分解成两个不相关的矩阵，而是X=A(A)^T

1. **No Spurious Local Minima in Nonconvex Low Rank Problems: A Unified Geometric Analysis**

证明在矩阵分解问题中的三种情况：1.鞍点2.局部最优3.全局最优，最后都会归为一种情况：得到的平稳点就是全局最优。因为：1.没有高阶鞍点存在2.所有局部最优相互等价，即均为全局最优。

1. **Stochastic Gradient Descent Tricks（节选）**

作者：Leon Bottou