

Abstract

Low-rank matrix factorizations (LRMFs) are a class of linear models widely used in various fields such as machine learning, signal processing, and data analysis. These models approximate a matrix as the product of two smaller matrices, where the left matrix captures latent features—the most important components of the data—while the right matrix linearly decomposes the data based on these features. There are many ways to define what makes a component "important." Standard LRMFs, such as the truncated singular value decomposition, focus on minimizing the distance between the original matrix and its low-rank approximation. In this thesis, the notion of "importance" is closely linked to interpretability and uniqueness, which are key to obtaining reliable and meaningful results.

This thesis thus focuses on volume-based constraints and regularizations designed to enhance interpretability and uniqueness. We first introduce two new volume-constrained LRMFs designed to enhance these properties. The first assumes that data points are naturally bounded (e.g., movie ratings between 1 and 5 stars) and can be explained by convex combinations of features within the same bounds, allowing them to be interpreted in the same way as the data. The second model is more general, constraining the factors to belong to convex polytopes. Then, two variants of volume-regularized LRMFs are proposed. The first minimizes the volume of the latent features, encouraging them to cluster closely together, while the second maximizes the volume of the decompositions, promoting sparse representations. Across all these models, uniqueness is achieved under the core principle that the factors must be "sufficiently scattered" within their respective feasible sets.

Motivated by applications such as blind source separation (e.g., hyperspectral unmixing) and missing data imputation (e.g., in recommender systems), this thesis also proposes efficient algorithms that make these models scalable and practical for real-world applications.

Résumé

Les factorisations matricielles de faible rang (LRMFs) sont des modèles linéaires largement utilisés dans des domaines tels que l'apprentissage automatique, le traitement du signal et l'analyse de données. Ces modèles approchent une matrice en la décomposant en produit de deux matrices plus petites : la première capture les caractéristiques latentes, c'est-à-dire les composantes les plus importantes des données, tandis que la seconde décompose linéairement les données à partir de ces caractéristiques. Il existe cependant de nombreuses manières de définir ce qui rend une composante "importante". Les LRMFs classiques, comme la décomposition en valeurs singulières tronquée, se concentrent sur la minimisation de la distance entre la matrice originale et son approximation de faible rang. Dans cette thèse, l'importance d'une composante est étroitement déterminée par l'interprétabilité et l'unicité, des notions clés pour obtenir des résultats fiables et pertinents.

Cette thèse explore donc des contraintes et régularisations volumiques visant à renforcer l'interprétabilité et l'unicité. Dans un premier temps, nous introduisons deux nouvelles variantes de LRMFs à contraintes volumiques. La première suppose que les points du jeu de données sont naturellement bornés (ex: des films notés entre 1 et 5 étoiles) et peuvent être expliqués par des combinaisons convexes de caractéristiques bornées de la même manière, permettant ainsi de les interpréter comme les données. Le second modèle est plus général et contraint les facteurs à appartenir à des polytopes convexes. Par ailleurs, nous proposons deux variantes de LRMF avec régularisation volumique : la première minimise le volume des caractéristiques latentes, favorisant ainsi un rapprochement entre elles, tandis que la seconde maximise le volume des décompositions, encourageant des représentations parcimonieuses. Dans l'ensemble de ces modèles, l'unicité est assurée par le principe clé selon lequel les facteurs doivent être "suffisamment dispersés" dans leur ensemble de solutions possibles.

Motivée par des applications telles que la séparation de sources aveugles (par exemple, le démélange hyperspectral) et l'imputation de données manquantes (par exemple, dans les systèmes de recommandation), cette thèse propose également des algorithmes efficaces permettant à ces modèles d'être adaptés à des applications réelles.