

# Métodos de Otimização Contínua para o Problema de Fatoração de Matrizes Não-Negativas

**Amanda T. Zanette\***

**Douglas S. Gonçalves†**

Departamento de Matemática, Universidade Federal de Santa Catarina,

88040-900, Campus Trindade, Florianópolis, SC

E-mail: amandatzanette@gmail.com.br

## RESUMO

A Fatoração de Matrizes Não-Negativas (NMF, do inglês *Non-negative Matrix Factorization*) consiste em decompor uma matriz  $X \in \mathbb{R}_+^{m \times n}$  como o produto de duas matrizes não-negativas  $W \in \mathbb{R}_+^{m \times r}$  e  $H \in \mathbb{R}_+^{r \times n}$ , de forma que  $X \approx WH$  [5].

Matematicamente, o problema pode ser formulado como

$$\min_{W, H} F(W, H) := \frac{1}{2} \|X - WH\|_F^2 \quad \text{sujeito a } W \geq 0, H \geq 0, \quad (1)$$

em que  $W \geq 0$  significa que cada entrada da matriz  $W$  deve ser não-negativa.

Trata-se de um problema de otimização não convexo, porém convexo em cada variável quando a outra é mantida fixa, o que naturalmente sugere o uso de métodos de minimização alternada [2]. A NMF possui aplicação em diversas áreas, incluindo compressão e segmentação de imagens, agrupamento de documentos e descoberta de padrões moleculares [6].

Neste trabalho, estudamos dois métodos clássicos para a resolução do problema (1): o método das Atualizações Multiplicativas e o método do Gradiente Projetado. O primeiro, proposto por Lee e Seung [5], pode ser interpretado como um método de minimização alternada que tenta minimizar uma função auxiliar  $G(W, H)$  que majoriza a função objetivo  $F(W, H)$ . A cada iteração, minimiza-se  $G$  com relação a uma variável mantendo a outra fixa, garantindo que o valor de  $F$  não aumente. As regras de atualização, componente-a-componente, são dadas por

$$W_{ia} \leftarrow W_{ia} \frac{(XH^T)_{ia}}{(WHH^T)_{ia}}, \quad H_{bj} \leftarrow H_{bj} \frac{(W^T X)_{bj}}{(W^T WH)_{bj}}, \quad (2)$$

que preservam automaticamente a não-negatividade – desde que o ponto inicial seja estritamente positivo – e asseguram o decaimento monótono da função objetivo. Apesar de sua simplicidade e estabilidade numérica, o método não garante convergência para pontos estacionários [4].

Já o método do Gradiente Projetado, estudado por Lin [6], aborda o problema por meio de minimizações alternadas em cada bloco de variáveis, resolvendo sequencialmente os subproblemas convexos

$$\min_{W \geq 0} F(W, H^k) \quad \text{e} \quad \min_{H \geq 0} F(W^{k+1}, H). \quad (3)$$

Cada iteração interna (no subproblema) realiza uma descida na direção do gradiente seguida de projeção sobre o ortante positivo. Para cada bloco de variáveis, as atualizações são dadas por

$$W^{(\ell+1)} = [W^{(\ell)} - \alpha_\ell \nabla_W F(W^{(\ell)}, H^k)]_+, \quad H^{(\ell+1)} = [H^{(\ell)} - \beta_\ell \nabla_H F(W^{k+1}, H^{(\ell)})]_+, \quad (4)$$

---

\*Graduanda em Matemática, Universidade Federal de Santa Catarina.

†Orientador, Universidade Federal de Santa Catarina.

em que  $[\cdot]_+$  denota a projeção sobre o conjunto das matrizes não negativas, isto é,  $[A]_+ = \max(A, 0)$  elemento a elemento. Os parâmetros  $\alpha_\ell$  e  $\beta_\ell$  são determinados por regras de passo adaptativas, como a de Barzilai–Borwein [1], de modo a garantir estabilidade e decaimento da função objetivo.

Sob condições apropriadas, é possível provar que todo ponto limite da sequência gerada é estacionário para o problema (1). A análise teórica segue o arcabouço de métodos de primeira ordem apresentado por Beck [2] e Bertsekas [3], que estabelecem propriedades de convergência global e estratégias eficientes para a escolha do tamanho de passo.

Os algoritmos estão sendo implementados em Julia, utilizando inicializações estritamente positivas e critério de parada baseado na variação relativa entre iterações consecutivas. Os experimentos numéricos avaliarão o desempenho dos métodos em termos de velocidade de convergência e qualidade da fatoração obtida. Espera-se que os resultados evidenciem a relação entre o comportamento numérico dos métodos e suas propriedades teóricas de convergência.

**Palavras-chave:** *fatoração não-negativa de matrizes, otimização, gradiente projetado, atualizações multiplicativas.*

## Referências

- [1] J. Barzilai e J. M. Borwein, Two-Point Step Size Gradient Methods, *IMA Journal of Numerical Analysis*, 8 (1988) 141-148.
- [2] A. Beck, “First-Order Methods in Optimization”, SIAM, 2017.
- [3] D. P. Bertsekas, “Nonlinear Programming”, 2nd ed., Athena Scientific, 1999.
- [4] E. F. Gonzales e Y. Zhang, *Accelerating the Lee-Seung algorithm for non-negative matrix factorization*. Technical report, Department of Computational and Applied Mathematics, Rice University, 2005.
- [5] D. D. Lee and H. S. Seung, Algorithms for Non-negative Matrix Factorization, *Advances in Neural Information Processing Systems*, 13 (2001) 556–562.
- [6] C.-J. Lin, Projected Gradient Methods for Non-negative Matrix Factorization, *Neural Computation*, 19 (2007) 2756-2779.