

Considere a seguinte matriz com avaliações de filmes feitas por usuários de um serviço de *streaming*:

Filme/Usuário	Carol	Lorayne	Wilian	Manu	Thayse
The Notebook	4	4	?	?	3.5
Whiplash	4	4.5	?	?	4.5
Os Oito Odiados	?	?	4.5	?	5
Os Miseráveis	2.5	?	?	5	?
Rocky I	1	?	3.5	4.5	5
Toy Story 3	3.5	5	3	5	5

O objetivo de um [Sistema de Recomendação de Filmes](#) é prever a nota que um usuário daria para um filme que ele ainda não assistiu. Se a nota prevista for suficientemente alta, o filme é recomendado.

Suponha que os filmes considerados são caracterizados apenas pelo nível de ação e de romance, de acordo com a tabela abaixo:

Filme	Ação f_1	Romance f_2
The Notebook	0.1	1.0
Whiplash	0.8	0.2
Os Oito Odiados	1.0	0.0
Os Miseráveis	0.6	0.4
Rocky I	1.0	0.2
Toy Story 3	0.9	0.1

QUESTÕES:

- (a) **Para cada usuário**, estime as pontuações y dos filmes não assistidos com base no modelo de regressão linear

$$y = \theta_0 + \theta_1 f_1 + \theta_2 f_2.$$

Descreva o problema de mínimos quadrados linear correspondente, utilizando a sua linguagem de programação preferida. Obtenha os coeficientes $\theta_0, \theta_1, \theta_2$ e apresente os códigos usados.

- (b) Para cada usuário, recomende os filmes cujas notas previstas pelo modelo sejam ≥ 4.5 .

OBS: Explique cada passo da sua resolução, seguindo o que foi dado em sala de aula ou baseado nas duas referências principais do curso.