

Questão 1

Coefficiente de regularização utilizado (taxa de acertos):

675.58805031572205734846647829

Coefficiente de regularização para erro quadrático: 55.71523605095193687475330080

A matriz confusão obtida pelo classificador na questão 1 está expressa abaixo:

2119,	12,	3,	3,	7,	0,	9,	25,	1,	0
94,	1628,	40,	33,	6,	76,	44,	56,	5,	33
41,	61,	1739,	6,	37,	25,	35,	46,	44,	16
32,	5,	3,	1744,	14,	18,	5,	18,	91,	3
27,	14,	160,	34,	1313,	65,	23,	79,	54,	48
33,	19,	1,	15,	12,	1821,	0,	14,	1,	27
76,	11,	20,	41,	3,	0,	1862,	3,	96,	28
178,	25,	75,	28,	64,	22,	12,	1507,	57,	23
19,	7,	39,	118,	5,	1,	132,	12,	1581,	19
2,	6,	5,	11,	19,	22,	4,	27,	1,	1902

Tabela 1: Matriz de confusão do classificador linear.

Cada linha da matriz representa, respectivamente, de 1 a 10, cada classe pretendida pelo treinamento (onde a classe 10 representa o dígito zero) e, nas colunas, quantas vezes cada classe, também de 1 a 10, foi escolhida pelo classificador quando este deveria escolher aquela da referida linha.

Nota-se que a classe reconhecida com maior precisão pelo sistema após o treinamento foi o dígito 1, sendo que a classe que mais gerou equívocos foi a do dígito 5, muitas vezes confundida com o dígito 3.

Nas figuras 1 e 2, temos exemplos de dígitos que foram reconhecidos erroneamente pelo classificador. Na figura 1, percebe-se que a curva superior distorcida e a curva inferior exageradamente grande fizeram o dígito 3 ficar parecido com o 5, fazendo o classificador se confundir. Para a figura 2, a parte superior do 5 foi feita de maneira muito discreta, fazendo o classificador classificá-lo como um 3.

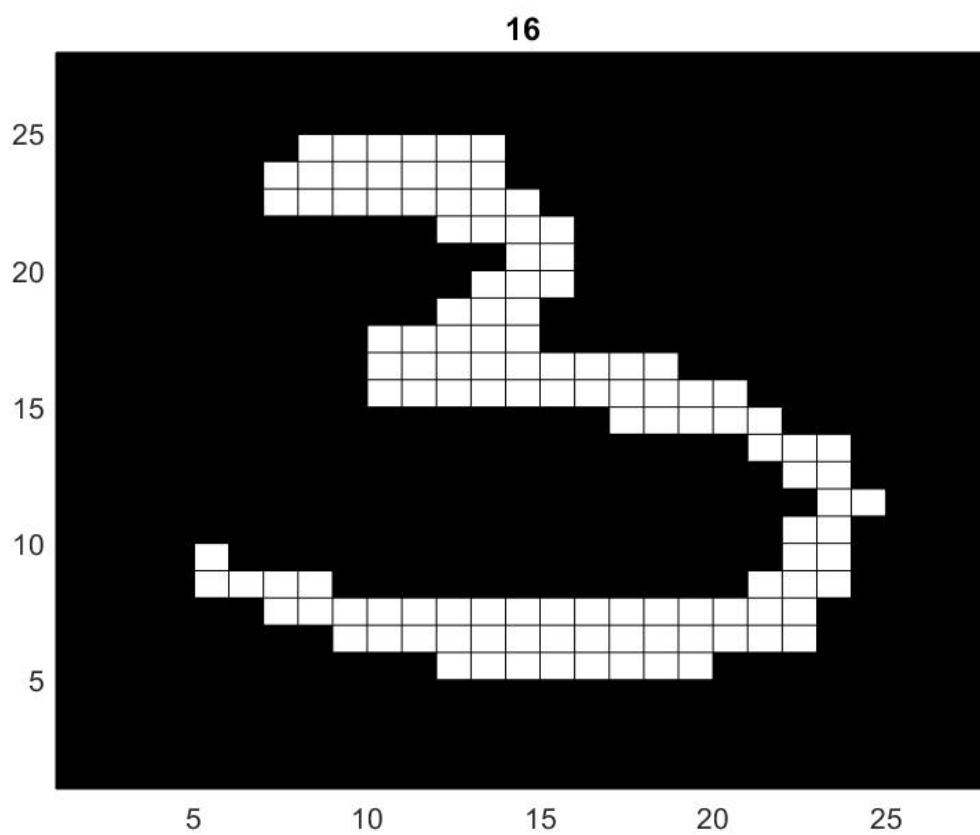


Figura 1: Um dígito três que foi reconhecido como cinco.

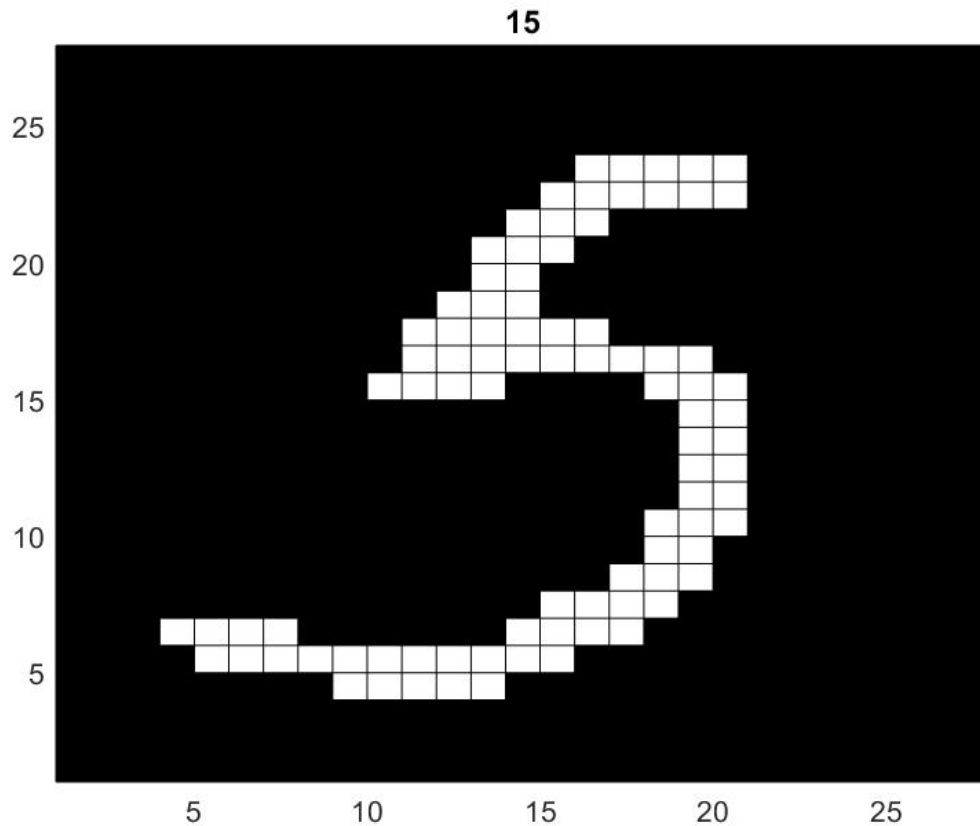


Figura 2: Um dígito cinco que foi reconhecido como três.

Gráficos de busca do coeficiente de regularização no anexo 1.

Gráficos de calor da matriz W no anexo 2.

Questão 2

Coeficiente de regularização utilizado (taxa de acertos):

64.000000000000000000000000000000

Coeficiente de regularização para erro quadrático: 24.25146506416636427161392930

Q2.1)

A matriz confusão obtida pela ELM na questão 2 está expressa abaixo:

2140,	11,	3,	3,	7,	6,	1,	5,	2,	1
18,	1824,	28,	21,	6,	27,	29,	32,	7,	23
11,	52,	1819,	2,	44,	17,	22,	49,	25,	9
20,	7,	0,	1780,	5,	14,	5,	8,	94,	0
19,	11,	76,	11,	1575,	49,	12,	14,	25,	25
4,	14,	0,	14,	17,	1860,	1,	12,	2,	19
34,	26,	6,	31,	6,	1,	1962,	3,	54,	17
45,	24,	58,	11,	47,	17,	17,	1714,	37,	21
9,	8,	26,	63,	12,	4,	63,	24,	1712,	12
1,	11,	3,	7,	11,	16,	1,	19,	3,	1927

Tabela 2: Matriz de confusão da ELM.

A matriz de confusão da ELM segue a mesma distribuição da questão 1. Nota-se que a taxa de acertos em todas as classes aumentou, mas o dígito mais bem reconhecido continua sendo o 1, provavelmente por seu formato simples e singular, enquanto que o dígito 5 continua sendo o mais complexo para a máquina, frequentemente confundido com o 3, provavelmente por seu traçado passar por muitos pontos em comum.

Na figura 3, foi reconstruído um dígito 5 que foi reconhecido como 3, provavelmente por seu formato distorcido. Na figura 4, um dígito 3 reconhecido como 5, provavelmente por suas curvas acentuadas e distorcidas.

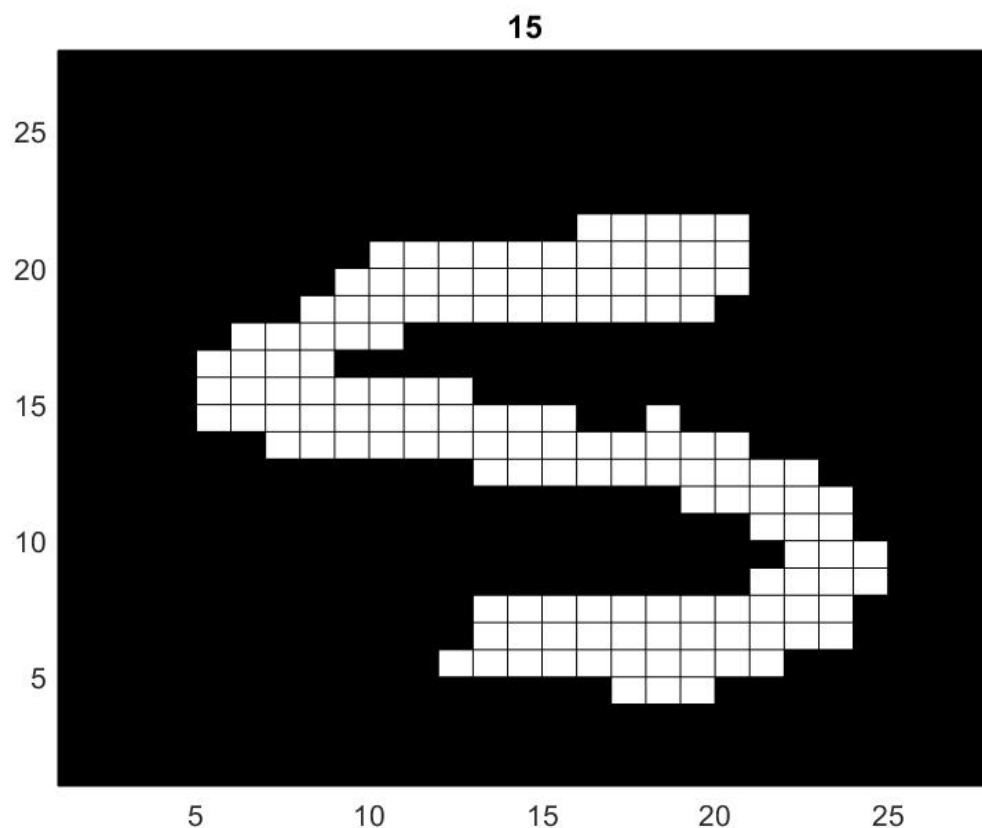


Figura 3: Um dígito cinco que foi reconhecido como três.

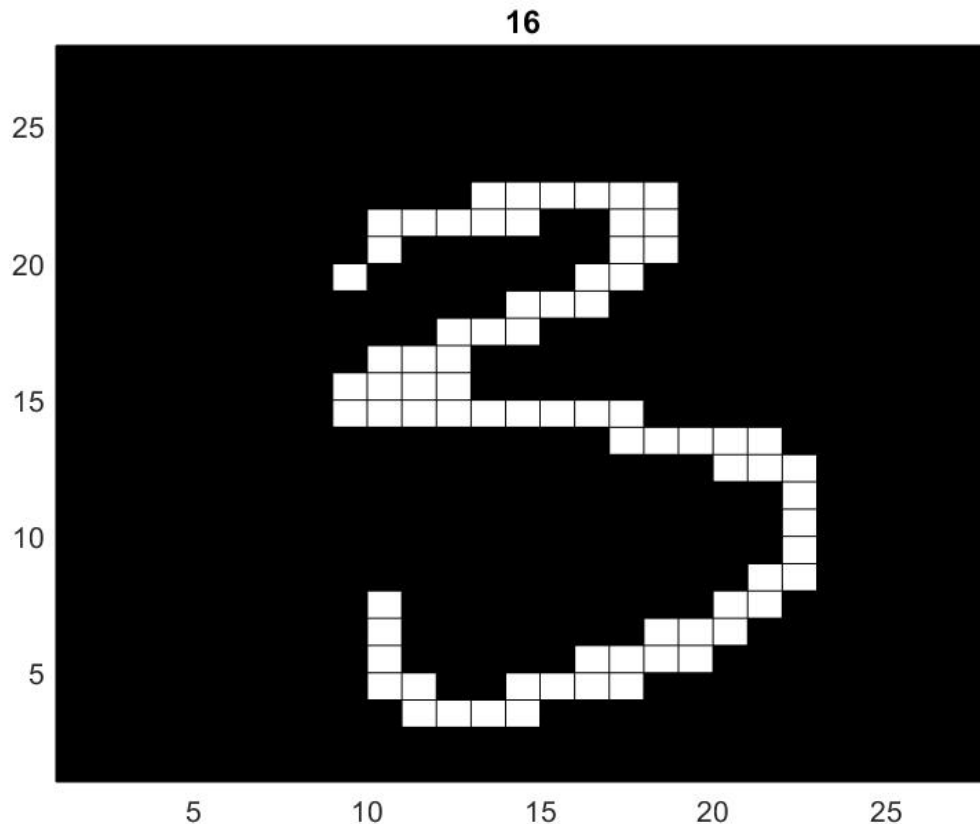


Figura 4: Um dígito três que foi reconhecido como cinco.

Q2.2) O ganho de desempenho se deve ao mapeamento com a função de tangente hiperbólica, que permite a saturação da saída do neurônio e evita que poucas entradas bem alinhadas, porém de magnitude alta, induzam um resultado incorreto. Assim, ficam sendo mais influentes as entradas com muitos pontos coincidentes aos maiores valores dos classificadores, em detrimento daquelas que possuem poucos pontos de grande valor sendo multiplicadas pelos elementos do respectivo classificador.

A diminuição do tempo de execução se deve ao fato de que a matriz de entrada e a matriz W reduziram de 784 para 500 (sem considerar o offset) seu número de colunas e linhas, respectivamente. Com isso, a quantidade de operações diminui significativamente, pois a quantidade de elementos a serem operados nestas matrizes é grande e seu número foi reduzido em pouco mais de 1 terço.

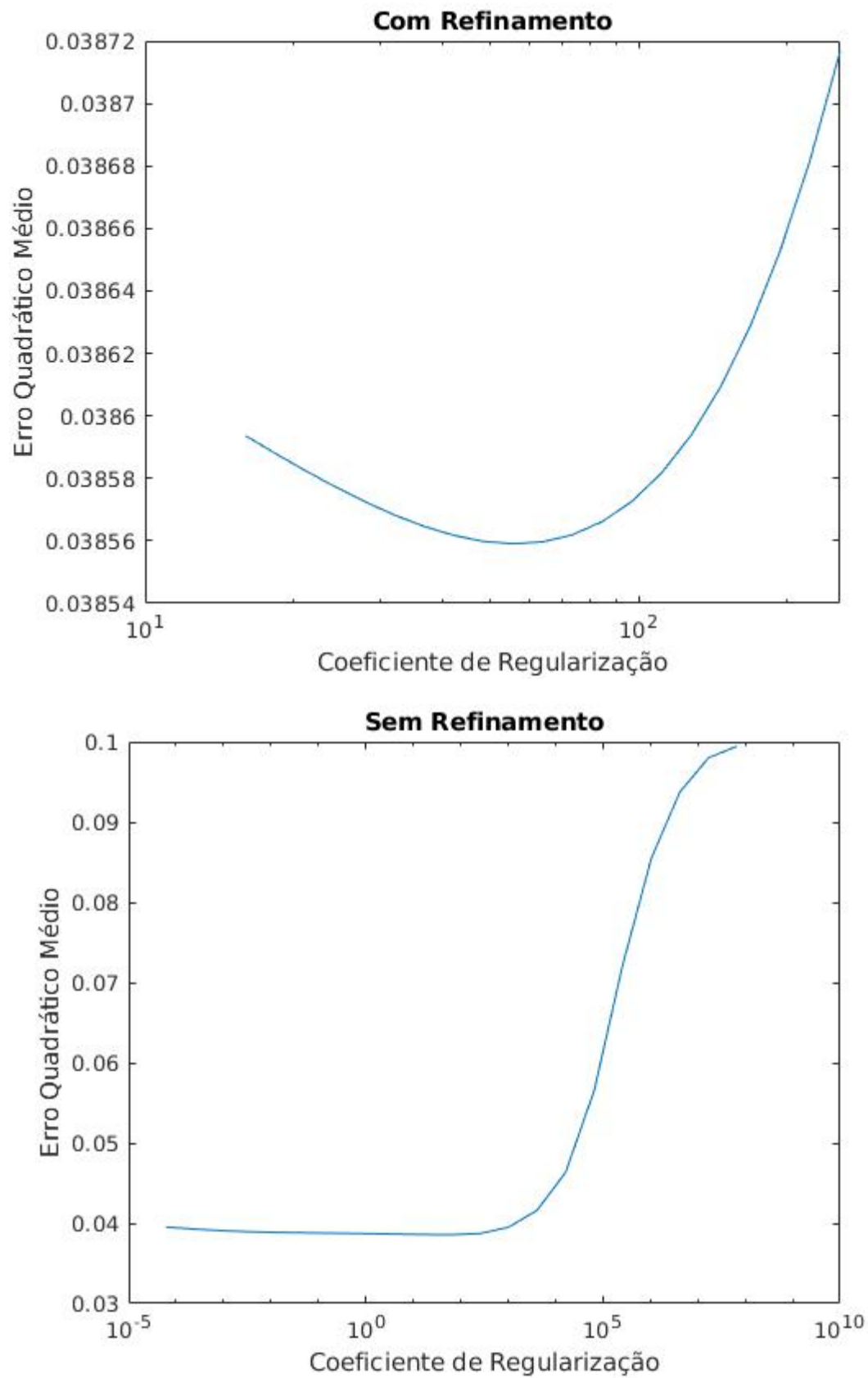
Q2.3) Como a matriz H é produzida através da multiplicação entre a entrada X e a matriz V de números normalmente distribuídos, o posto de H é maior que o de X , o que faz com que a necessidade de regularização no caso da ELM seja menor. Como calculamos a regularização de forma a penalizar o crescimento dos vetores de W , o coeficiente de regularização será o menor possível para realizar a tarefa. Portanto, o coeficiente de regularização utilizado no classificador linear (675.588) foi maior do que para a ELM (64.000).

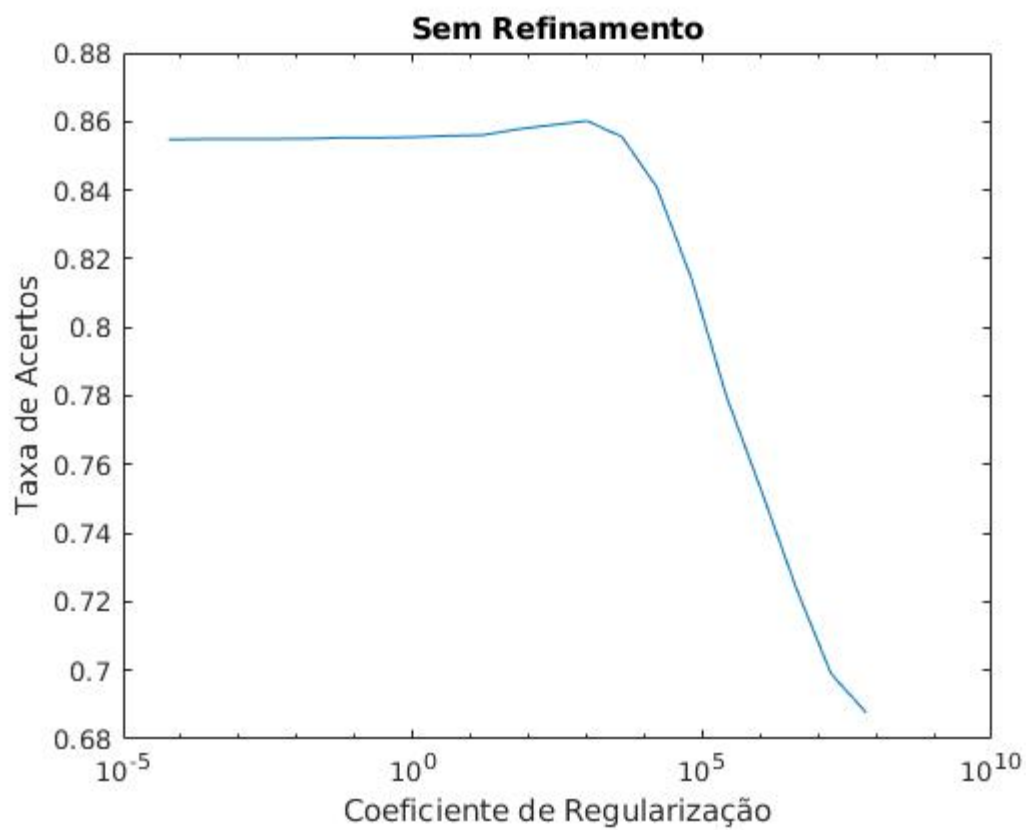
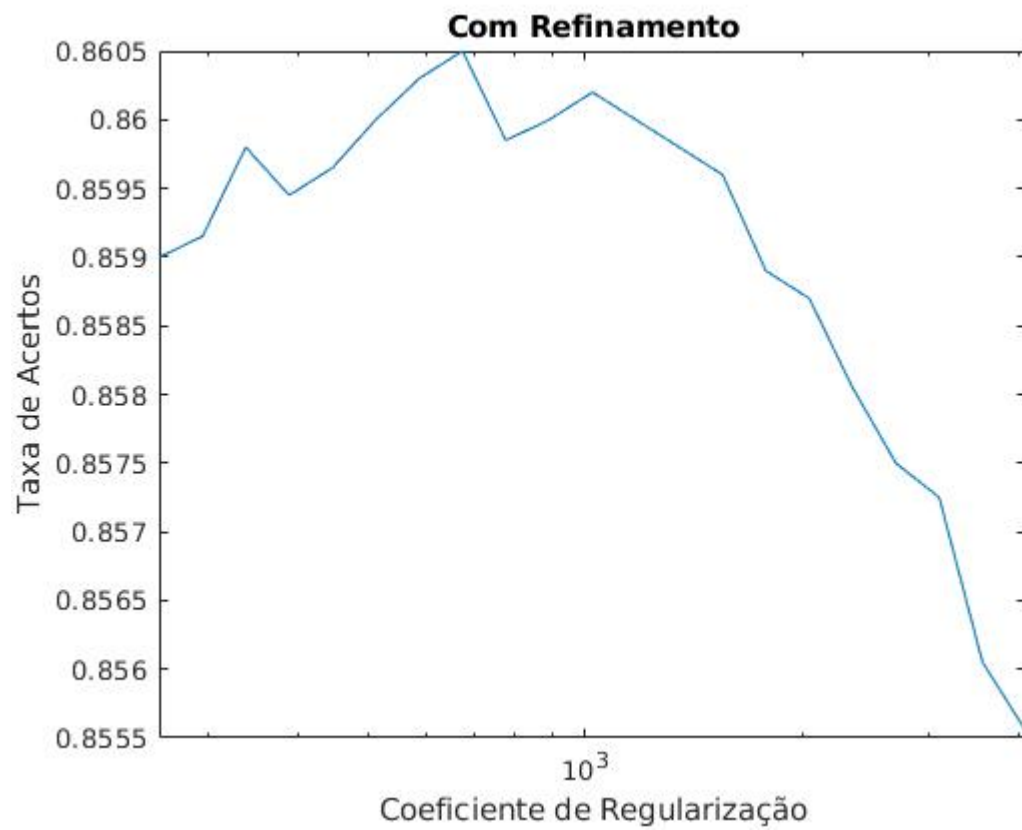
Q2.4) Como a matriz de entrada será diferente para cada caso dos pesos aleatórios, o

coeficiente de regularização pode ser alterado, para mais ou para menos, dentro de uma faixa de valores que é relativa ao desvio padrão escolhido para a distribuição normal dos pesos aleatórios.

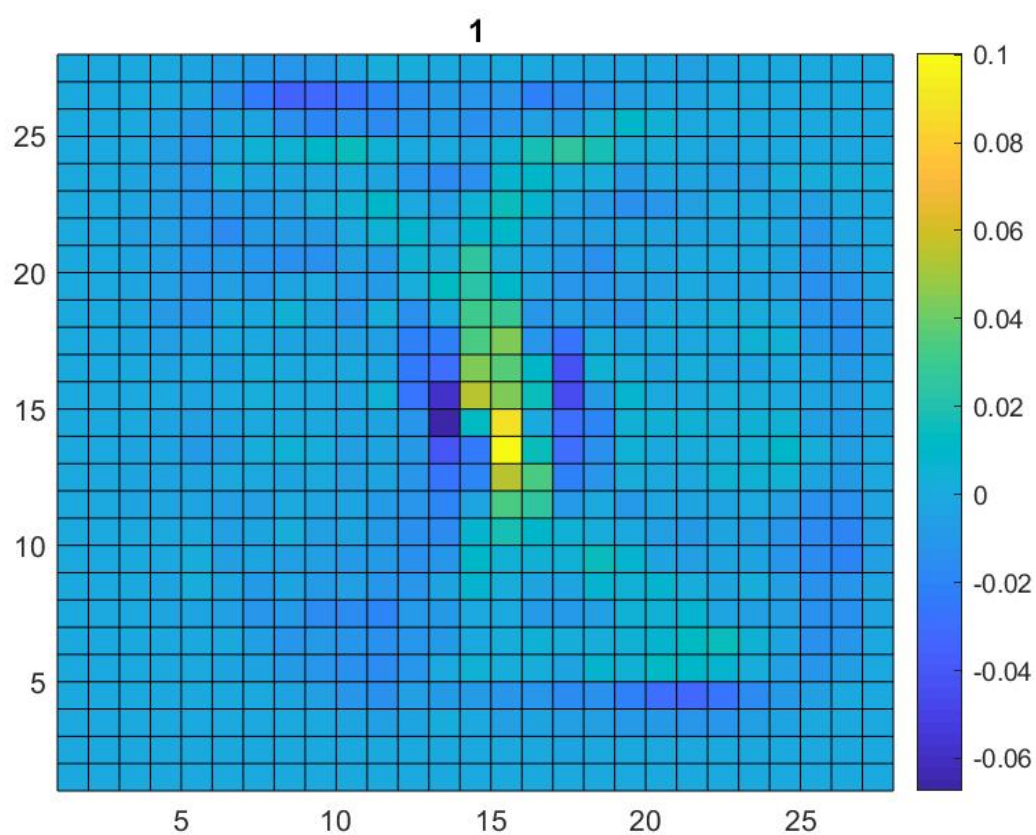
Gráficos de busca do coeficiente de regularização no anexo 3.

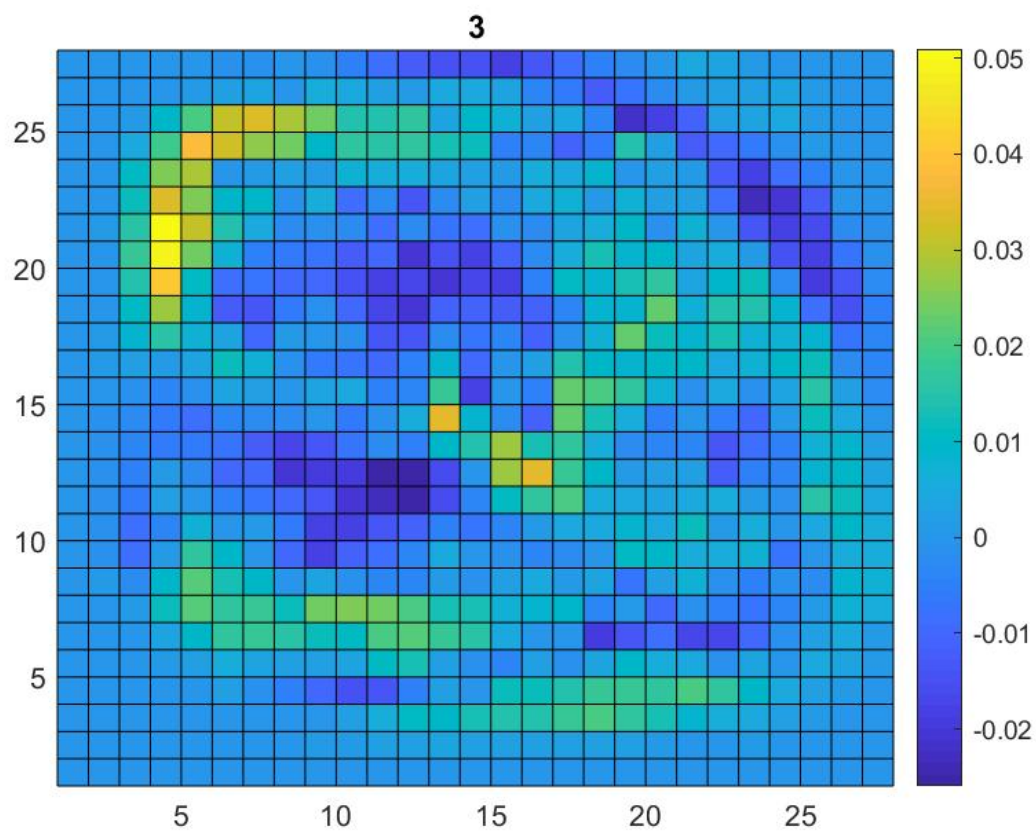
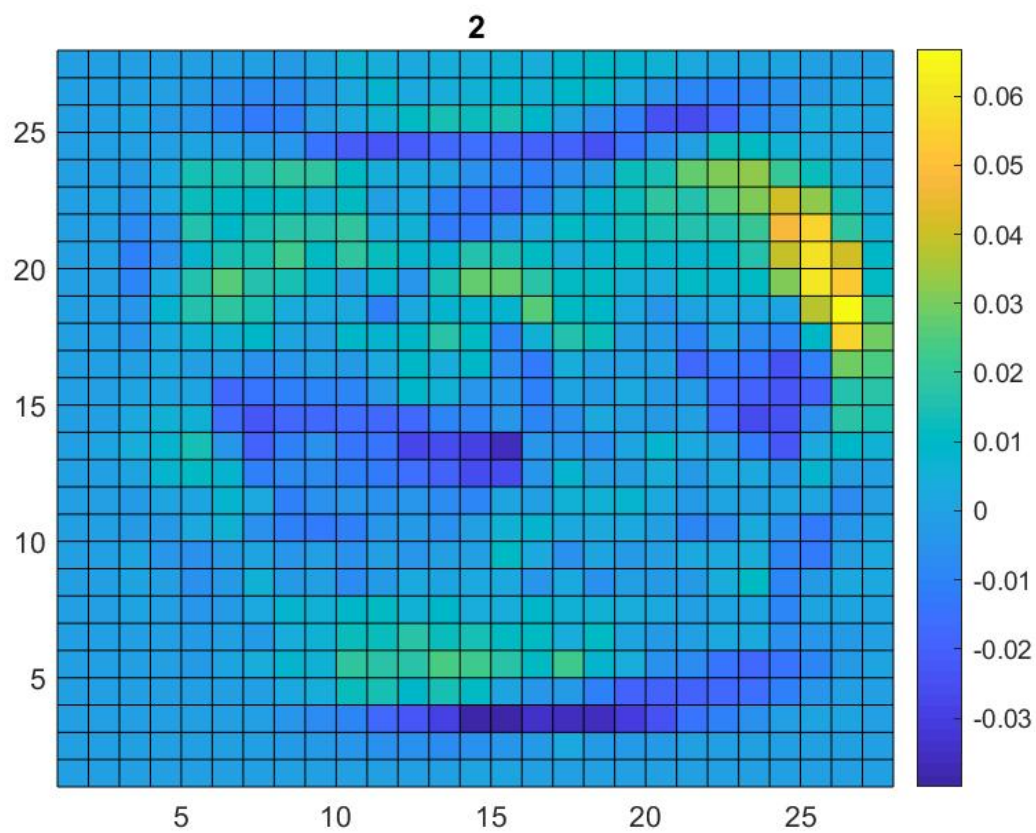
Anexo 1 - Busca pelo coeficiente de regularização da atividade 1

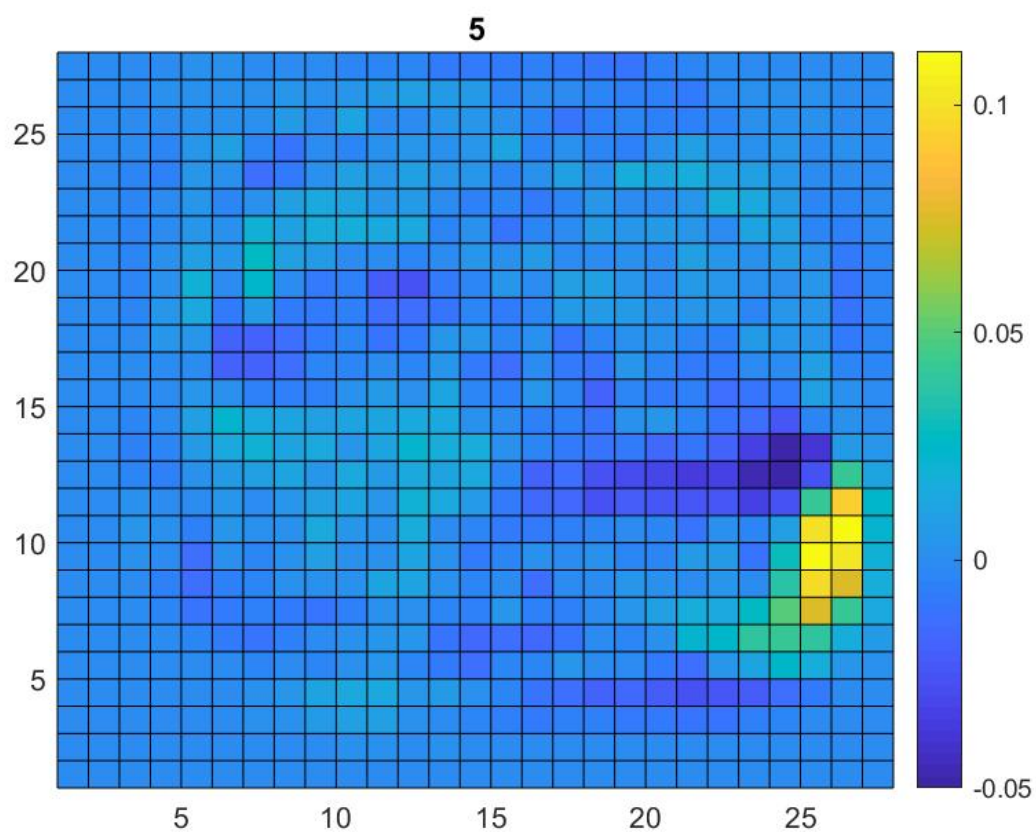
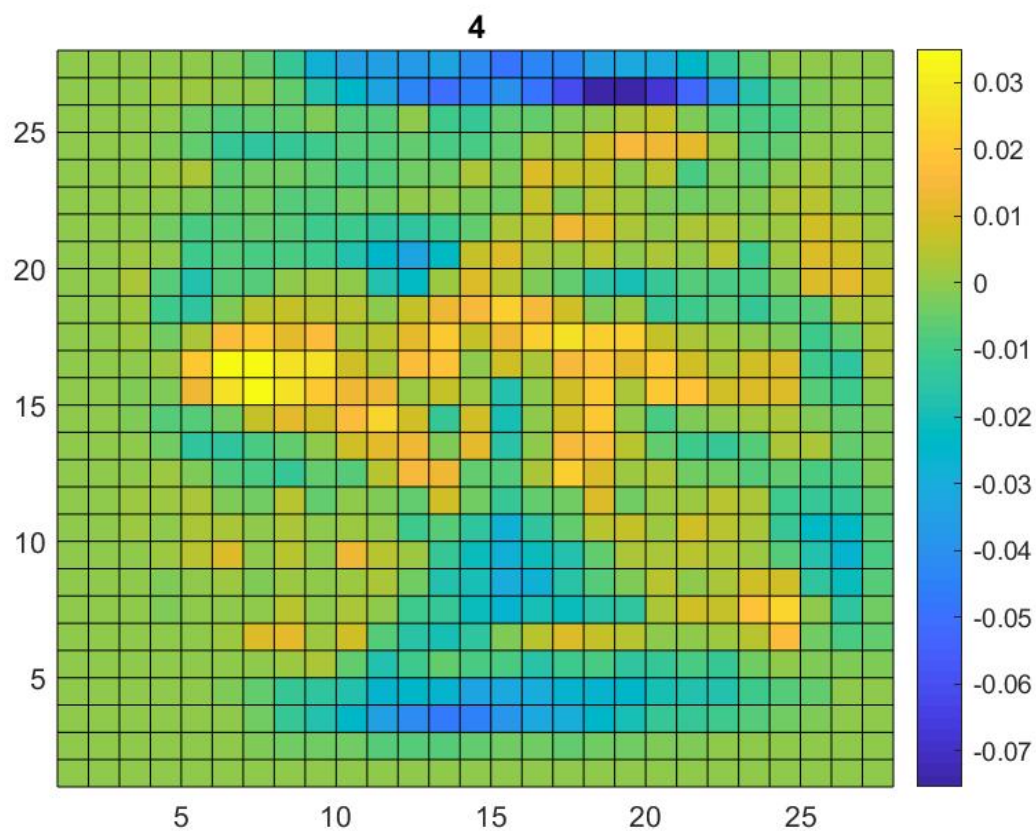


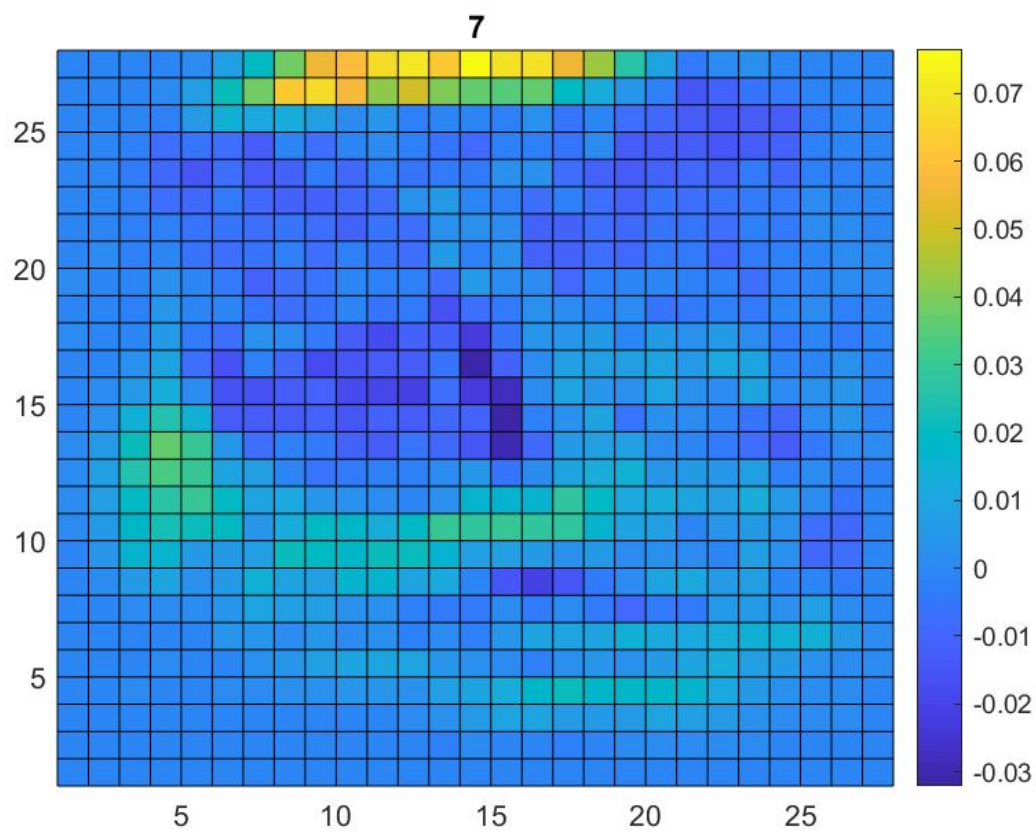
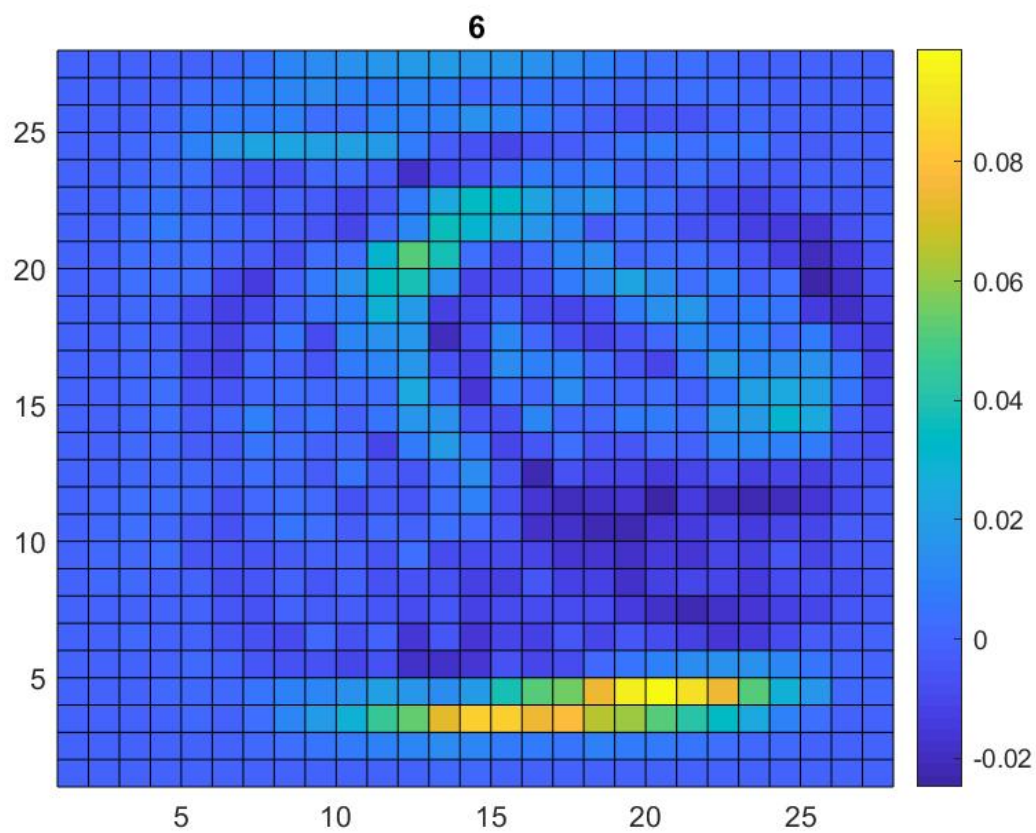


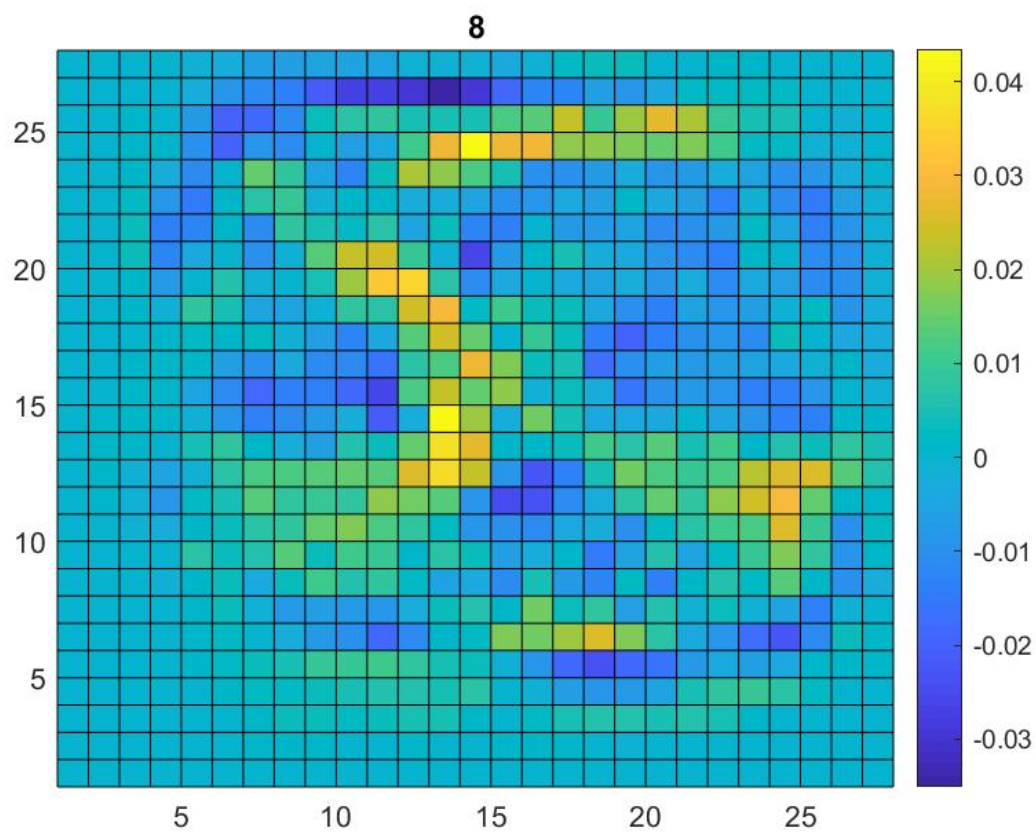
Anexo 2 - Gráficos de calor da atividade 1

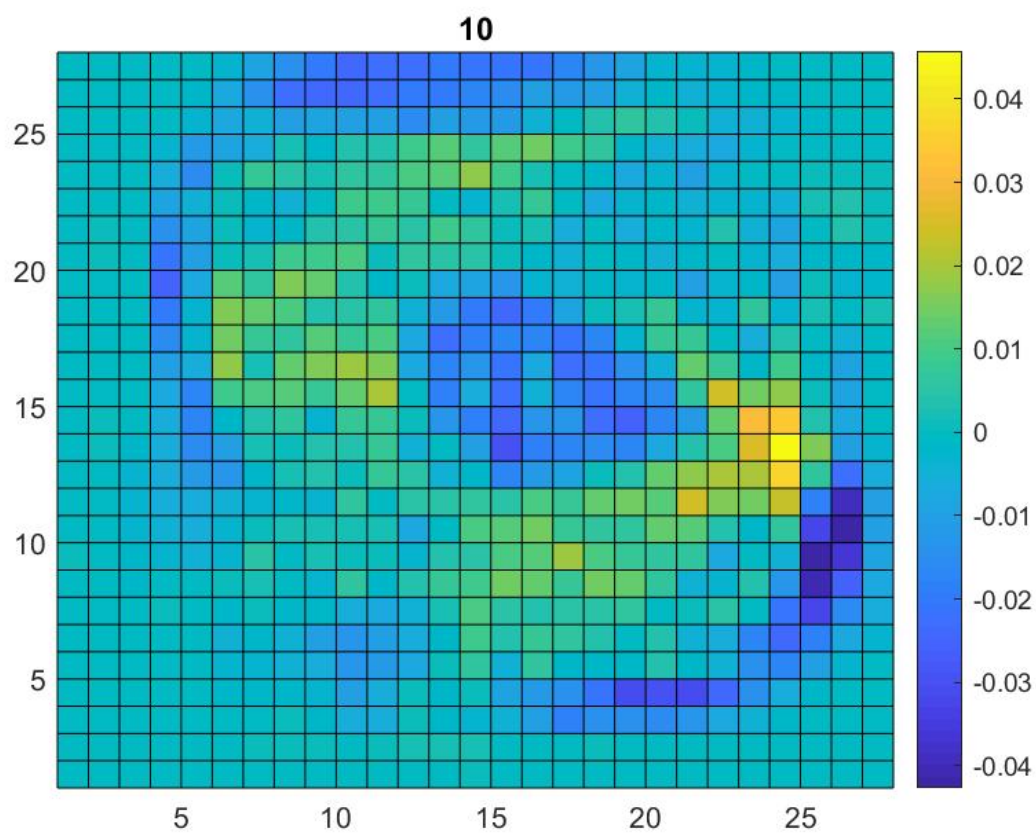
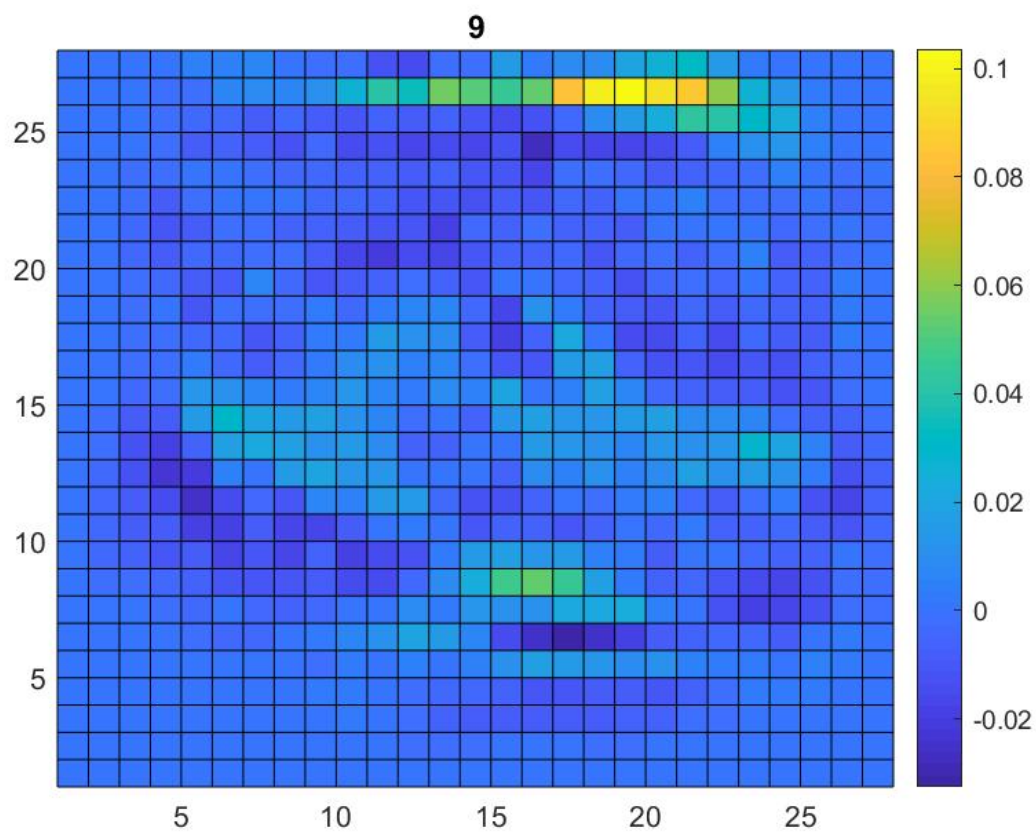












Anexo 3 - Busca pelo coeficiente de regularização da atividade 2

