

Projeto 2: Mínimos quadrados e otimização

O desafio:

- O objetivo é produzir um modelo preditivo $f^*(\mathbf{x})$ a partir de um conjunto de dados (\mathbf{X}, \mathbf{y}) .
- A metodologia será utilizando mínimos quadrados. Apenas poderá ser utilizada a função pronta de otimização `minimize`. Para usar outras consultar com os docentes.
- Cada grupo deverá também programar o método do gradiente "from scratch" e comparar com a função `minimize`. A comparação deverá fazer parte do relatório final.
- Procuram-se modelos preditivos com poucas variáveis, isto é, que usem poucas colunas da matriz de dados \mathbf{X} . Cada grupo há de determinar os melhores modelos possíveis para cada número R de colunas envolvidas, sendo $R = 1, 2, 3$ e 4 .
- Cada grupo receberá também um conjunto de teste $\mathbf{X}_{\text{teste}}$, para o qual será gerada uma previsão de resposta \mathbf{y}_{prev} . Os modelos propostos serão avaliados calculando $\|\mathbf{y}_{\text{prev}} - \mathbf{y}_{\text{teste}}\|$ (em média quadrática), sendo $\mathbf{y}_{\text{teste}}$ o vetor de respostas verdadeiras (mantido em segredo).
- Cada grupo deverá entregar um relatório explicando o desenvolvimento dos modelos propostos. Também deverá entregar um arquivo contendo os vetores \mathbf{y}_{prev} correspondentes a cada modelo (para cada valor de R).
- Cada grupo deverá também preparar uma apresentação de 5 minutos explicando como chegou-se aos modelos finalmente propostos.

Comentários e sugestões:

- Os docentes (Gustavo, Maria Victória, Felipe) fazemos parte dos grupos! Em particular, podemos ajudar nos períodos de monitoria, como Terças-feiras, das 16 às 18 hs, na sala G2 do STI-EESC (acima do Banco do Brasil, do lado do bloco C de aulas, frente à biblioteca da EESC).
- Encaminhar o resultado do projeto pelo Moodle até o dia 29/05. Pode ser um Colab Notebook ou outro formato, contendo relatório e códigos.
- Cada grupo deve reunir-se com um docente para mostrar o avanço do projeto até o dia 27/05 (pode ser em tempo de aula, de monitoria, ou marcado separadamente). Cada grupo é responsável de marcar e comparecer na reunião.