Aula 4: derivadas parciais e o Jacobiano

13/11/2026, on-line. O que não couber entra na aula seguinte.

1. Revisão das aulas 2 e 3

de memória: a camada como matriz, duas lineares, derivada parcial, gradiente
o fio do curso, da aula 1 à 10; hoje, a 4
$\partial f/\partial x_j$ com as outras paradas; ∇f aponta para onde f mais cresce; diferença central

2. O Jacobiano

e quando a saída é um vetor? duas entradas, duas saídas, quatro derivadas
J: uma linha por saída, uma coluna por entrada; $\Delta h \approx J\,\Delta x$; com m = 1, o gradiente
camada linear: J = W, em qualquer ponto, como a reta da aula 2
tanh de −1 a 1; $\tanh' = 1 - \tanh^2$, caixa-preta conferida pelo código
camada com tanh: linha k de W vezes $1 - h_k^2$; exemplo 0,5966 e 1; diferença $7 \cdot 10^{-11}$
vídeo: uma entrada de cada vez, para cima e para baixo; a variação sobre 2Δ é uma colu
a tabela: com Δ = 0,5, a coluna 2 sai 1,5232; com 0,001, 2,0000, a conta à mão
vídeo: o quadradinho e o paralelogramo de J, que quase coincidem de perto

3. A regra da cadeia com vetores

J de g ∘ f: o J de g, em f(x), vezes o J de f; a ordem importa; m × k vezes k × n
até a perda: $2(y - t) \cdot v \cdot J_2 \cdot J_1$ dá uma linha; L = 0,9766, gradiente (1,0940; −4,6074)
de trás para a frente: $3m^2$ contra $2m^3 + m^2$; m = 1.000, 3 milhões contra 2 bilhões
vídeo: as duas ordens em quadradinhos, com 4 neurônios: 48 contas contra 144
a retropropagação é esse produto, de trás para a frente: a aula 5

4. De onde vem

Leibniz, 1676; Jacobi, 1841; Linnainmaa, 1970; Baur e Strassen, 1983

5. O código

jacobiano.py só pela saída: é a referência da entrega
a entrega: o Jacobiano numérico de uma camada, comparado com o à mão

6. Fechamento

3Blue1Brown, redes neurais 3 e 4; Nielsen, capítulo 2
a seguir: aula 5, a retropropagação; aula 6, o MNIST em NumPy