

- b) All gradients get twice as long but after noma (izahm there i) he difference.
- C) Adding a constant does not affect gradients
- d) All gradients are flipped (v=-v)

 only this

 one is

 charged.

Li =
$$-\ln(P_1)$$
 =) $\frac{\partial L_i}{\partial w_1} = -\frac{1}{P_1} \cdot \frac{\partial P_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial w_1}$

$$\frac{\partial P_1}{\partial y_1} = P_1 (1-P_1), \quad \frac{\partial y_1}{\partial w_1} = \Gamma, \quad \frac{\partial y_1}{\partial c_1} = L$$

Update mle:
$$w_1 = w_1(u) + \mu(1-p_1) I$$

$$C_1(u+1) = C_1(u) + \mu(1-p_1)$$

$$L_i = -\ln(p_2) = \frac{\partial L_i}{\partial w_1} = -\frac{1}{p_2} \frac{\partial p_2}{\partial y_1} \frac{\partial y_1}{\partial w_1}$$

Update unle:
$$w_{i}^{(k)} = w_{i}^{(k)} - \mu P_{1} T$$

$$c_{1}^{(k+1)} = c_{1}^{(k)} - \mu P_{1}$$

$$26) \quad 10.(5.5.3+1) + 20(3.3.10+1) + 20(3.3.20+1) = 6200$$

- 3 b) See lab code.
 - b) The chance of picking five Miler will take wear is much smaller => It will take much longer to find a good solution.

- $\mathcal{L}_{\alpha} = \mathcal{L}_{\alpha} \quad \mathcal{L}_{\beta}$ $\mathcal{L}_{\alpha} = \mathcal{L}_{\alpha} \quad \mathcal{L}_{\beta}$
 - b) ...
 - C) Almost certainly wrong
 - b) 2 views => overditermined so pretty certain.
 - c) Vary good.

Time per iteration: 1+0.01.100 = 2

Rate of outlier-free subsets that

Alba More

O.13

Time per iteration. 0.01+0.01.100=1.01

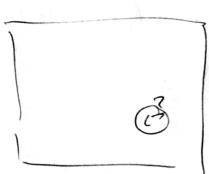
Rate of ontlin-free subsets: 0.14

Model 2 is 1.01/2 2 5 thus slower.

b) Change of variables \Rightarrow edge punt at (0,0) $M(x,y) \approx M(0,0) + (x,y) \not\vdash M + \frac{1}{2}(x,y) + (\frac{x}{y})$ Let $g = \binom{1/x}{1/y} = \forall T$. Set $\binom{x}{y} = s \cdot g$ We get a 1D tyler expansion perp

to the edge. Solve for a local max.

6a)



Prob of a vandom Mire: $\frac{\pi \cdot 2^2}{1000^7} \approx 10^{-5}$ 97 & 100 candidates => $n \approx 100$, $np \approx 10^{-3} = 2$ We get 3 inlies for free and $P(x=5) = \frac{(10^{-3})^5}{5!} = \frac{(-0.00)^{-3}}{5!} \approx 10^{-17}$

b) Another binarial distributed P(x > 5) is unegligible compared to P(x = 5) = 1 binomial distribute N = 1000 $p = 10^{-17} \Rightarrow \lambda = 10^{-14}$ $P(x = 1) \approx 10^{-14} e^{-(10^{-17})} \approx 10^{-14}$ P(x > 1) is negligible

Ans: 10^{-14}