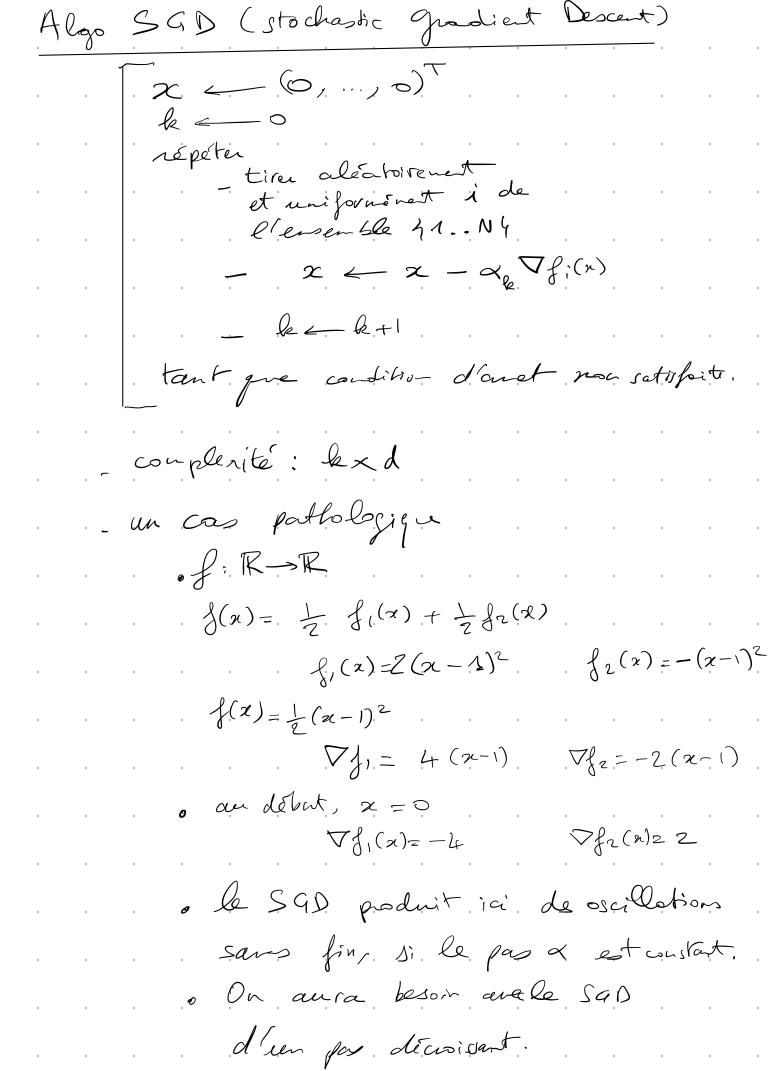


- ple avec la decente de gradient? - chaque itération reprient de posser. par tonts le dancées. - matrice A ERMXd (n exert d dinesons) Conplexité: lex mx d (où le= ub d'itératio) => or va s'interesse à des algos. moins gourmants en doncées par itération Ide: le ple d'agian qui nous intéressent (sur de dounts), sont souvent de la forme min f(x) où $f(x) = \frac{1}{N} \stackrel{\text{P}}{\leq} f(x)$

ex: régression livéaire. $f(x) = \prod_{i=1}^{N} f_i(x) \quad \text{où } f_i(x) = (a_i^* n - b_i)^2$ ex: régression logistique
- réseaux de neuvous.



Pourquoi la SGD fonctioner?

.

- Rappel sur l'espérance

si une variable alsatoire Z prend de valeurs Z. .. ZN avec probabilité li .- PN, alors on défini l'espérance de 2 E[Z]= Niz;

s; Zet le Zi sont le vectour. la définition et la mêne.

- Alétape le, quel et le vecteur The esperat?