Stochastic Approxi.

It's a method to solve $\begin{cases} \text{i) Solve equations} \\ \text{ii) Optimize Func's} \end{cases}$

We consider i):

To solve $f(x)=0$ without knowing $f(x)$. only given randomly noisy observation of $f(x)$.

① Model:

Given X_n. Observe $Y_n = f(X_n) + \eta_n$. where η_n i.i.d. $E(\eta_n)=0$. $Var(\eta_n)=1$. random noise. f is locally monotone $\uparrow$ (i.e. for x_0. $f(x_0)=0$. $\exists\,(x,\eta)$. s.t. $x_0 \in (x,\eta)$. WLOG. $f\uparrow$ on the interval (x,η). x_0 is its unique zero)

1°) Given previous observation: Y_n Choose X_{n+1}:

$X_{n+1} = X_n - a_n Y_n$. (a_n) Seq of suitable positive number

Rmk: It's reasonable to guess x_0 by its local monotone property.

2°) For (X_n) converges. We require $a_n Y_n \to 0$ $(n\to\infty)$

Note $Y_n \neq 0$. So we need: $a_n \xrightarrow{n\to\infty} 0$

Rmk: i) $a_n \to 0$ can't be too rapid which will moves x to x_0 by large distance

ii) $a_n \to 0$ should be rapid enough to damp out the noise.

② Thm: $f: \mathbb{R}^1 \to \mathbb{R}^1$. $X_0 \in L^2$. Consider (X_n): $Y_n = f(X_n) + \eta_n$.

$X_{n+1} = X_n - a_n Y_n$. Assumpt that:

i) X_0 indept with (η_n). i.i.d seq. $E(\eta_n) = 0$. $Var(\eta_n) = 1$.

ii) $\exists\, c \in (1, \infty)$. $|f(x)| \le c|x|$. $\forall x \in \mathbb{R}^1$.

iii) $\forall \delta > 0$. $\inf_{|x| > \delta} (x f(x)) = \varepsilon > 0$

iv) $a_n \geq 0$. $\sum a_n = \infty$. v) $\sum a_n^2 < \infty$. Then: $X_n \xrightarrow{n \to \infty} 0$. a.s.

Pf: It suffices to prove: $X_n^2 \xrightarrow{n \to \infty} 0$. a.s. (Advantage: $X_n^2 > 0$)

1) (X_n^2) is supermart. w.r.t $\mathcal{G}_n = \sigma(X_1 \cdots X_n)$?

Calculate: $E(X_{n+1}^2 \mid \mathcal{G}_n) \overset{(*)}{\le} X_n^2 (1 + a_n^2 c^2) + a_n^2 c^2 - 2 a_n X_n f(X_n)$

$\le X_n^2 (1 + a_n^2 c^2) + a_n^2 c^2$

Set $W_n = (X_n^2 + 1) / \prod_1^{n-1} (1 + a_k^2 c^2)$ supermart. ≥ 0

2) Show: $\lim X_n^2 = 0$. a.s.

Denote: $\prod_0^{n-1} (1 + a_k^2 c^2) = b_n$. Then: from above $(*)$:

$E(W_{n+1} | \mathcal{G}_n) \le W_n - 2 a_n b_{n+1} X_n f(X_n)$ --- (Δ)

$\forall \delta > 0$. Set $B_m = \bigcap_{n \ge m} \{ |X_n| \ge \delta \}$. prove: $P(B_m) = 0$.

Note: $\sum a_n^2 < \infty \Rightarrow b_n$ converges $\Rightarrow X_n^2$ converges.

By iii): $E(X_n f(X_n)) \geq E(X_n f(X_n) I_{B_m}) \geq \varepsilon P(B_m)$. $n \geq m$.

Put in (Δ) by taking expectation: for $m \le n$.

$P(B_m) \le E(W_m) / 2\varepsilon \sum_m^{n-1} a_k b_{k+1} \xrightarrow{n \to \infty} 0$ ($\sum a_k b_{k+1} \to \infty$)

Rmk: ii). iii) Control properties of $f(x)$. iv). v) control order of (a_n).