Stochastic Approximation

Fundamentals of AI Reading Group

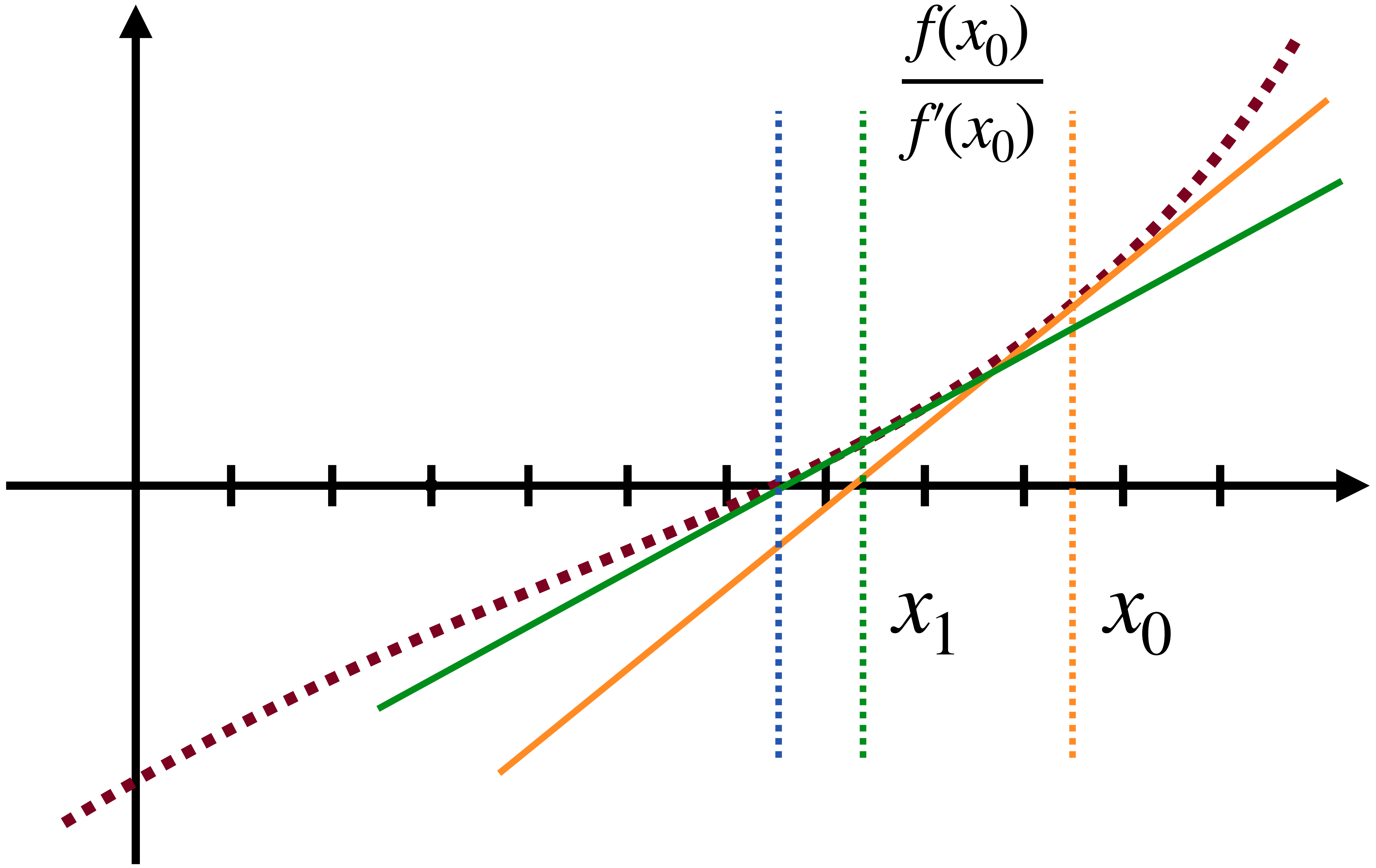
Paulo Oliva

Wednesday, 18 March 2026

Background
&
Motivation

The Newton-Raphson Method

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

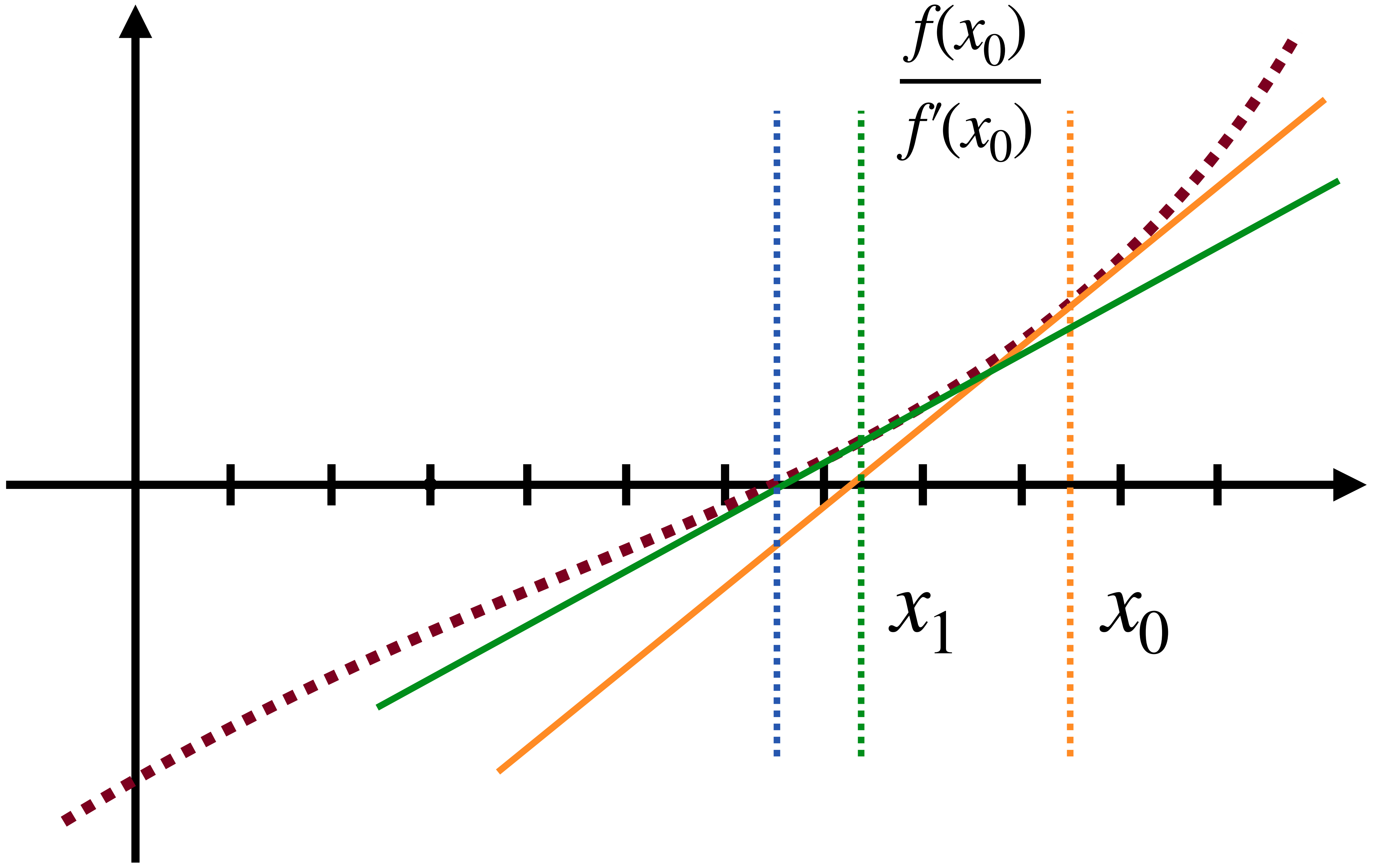


Stochastic Case

$$f(x) = \mathbb{E}(Y_x)$$

$$f(x), f'(x) = \text{unknown}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$



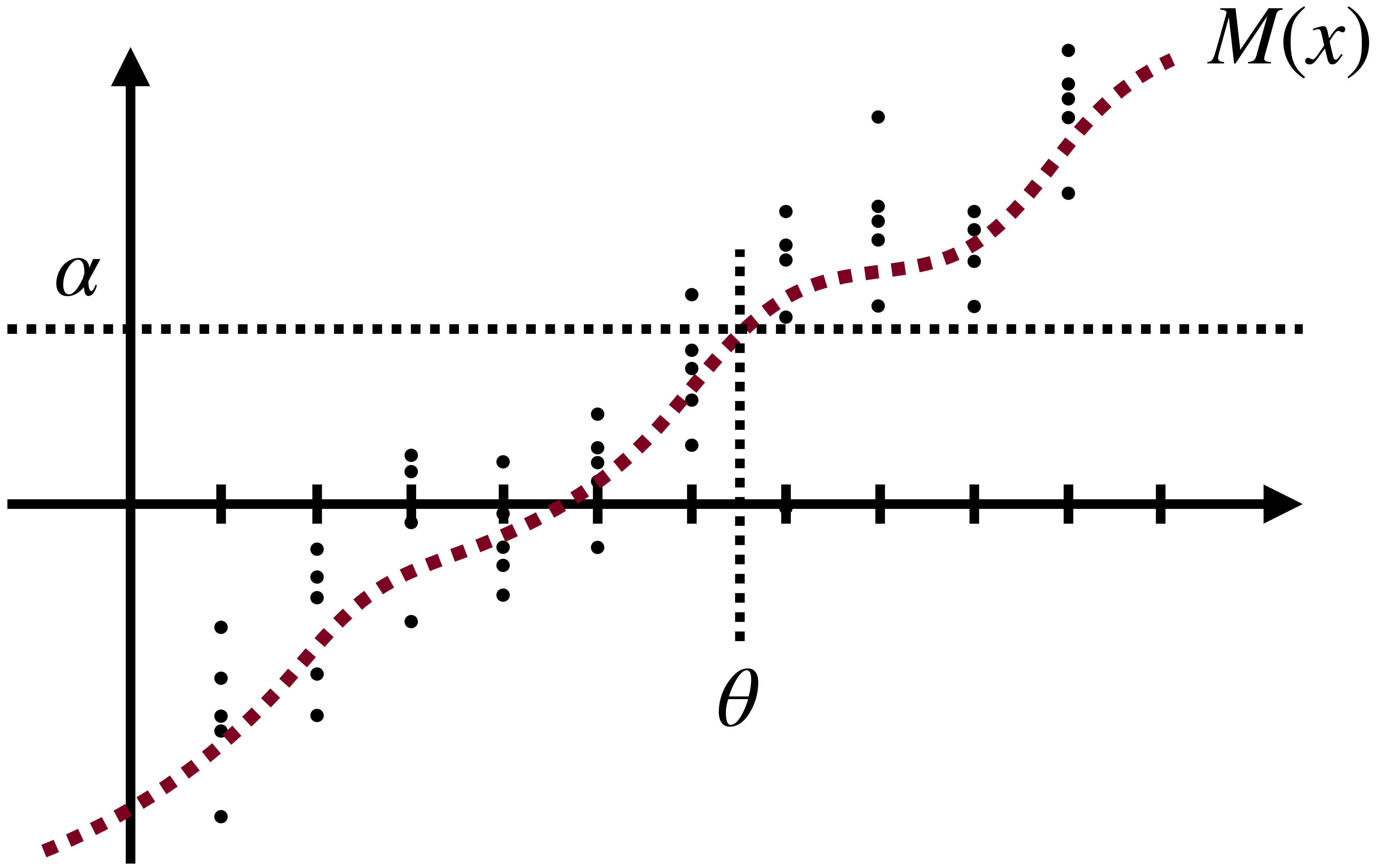
A STOCHASTIC APPROXIMATION METHOD[1]

By Herbert Robbins and Sutton Monro

University of North Carolina

1. Summary. Let $M(x)$ denote the expected value at level x of the response to a certain experiment. $M(x)$ is assumed to be a monotone function of x but is unknown to the experimenter, and it is desired to find the solution $x = \theta$ of the equation $M(x) = \alpha$, where α is a given constant. We give a method for making successive experiments at levels $x_1, x_2, \cdots$ in such a way that x_n will tend to θ in probability.

Ann. Math. Statist. 22(3): 400-407 (September, 1951). DOI: 10.1214/aoms/1177729586



- $M(x) = \mathbb{E}(Y_x)$
- A stage n, give sample of patients dose x_n
- Observe outcome $y_n = Y_{x_n} \in [0,1]$
- Update $x_{n+1} = x_n + a_n(0.5 - y_n)$

- $x =$ drug dose (response $\{0,1\}$)
- $\alpha = 0.5$
- $\theta =$ dose with 50 % success
- $Y_x \in [0,1]$

Convergence of Random Variables

A sequence of reals $(x_n)_{n \in \mathbb{N}}$ is said to **converge to** x if

$$\lim_{n \to \infty} |x_n - x| = 0$$

More explicitly

$$\forall \varepsilon > 0 \, \exists n \forall i \geq n \; (|x_i - x| < \varepsilon)$$

- For a sequence of **random variables** $(X_n)_{n \in \mathbb{N}}$ we have a few options:

(1) $\forall x \, (\lim_{n \to \infty} |\mathbb{P}(X_n \leq x) - \mathbb{P}(X \leq x)| = 0)$

easy

(2) $\forall \varepsilon > 0 (\lim_{n \to \infty} \mathbb{P}(|X_n - X| \leq \varepsilon) = 1)$

Fatou

(3) $\mathbb{P}(\lim_{n \to \infty} X_n = X) = 1$

Markov

(4) $\lim_{n \to \infty} \mathbb{E}(|X_n - X|^r) = 0$

convergence in distribution

convergence in probability

almost sure convergence

convergence in the r-mean

Robbins-Monro 1951 Seminal Paper

Robbins-Monro (1951)

- Y_x random variable parametrised by x
- $\mathbb{P}(Y_x)$ the probability of Y_x for given x
- Can only observe $\mathbb{P}(Y_x)$ via sampling
- $M(x) = \mathbb{E}(Y_x)$
- **Problem**:
  Find θ such that $M(\theta) = \alpha$

<u>Stochastic Approximation Algorithm</u>

$$X_{n+1} = X_n + a_n(\alpha - Y_n)$$

Robbins-Monro (1951):

X_n converges (2-mean) to (unique) solution θ

Assumes:

- $a_n \to 0,\ \Sigma\, a_n = \infty,\ \Sigma\, a_n^2 < \infty$
- Y_x bounded w. p. 1
- function $M(x) = \mathbb{E}(Y_x)$
  - non-decreasing
  - θ unique solution for $M(x) = \alpha$
  - derivative at solution is positive

Generalising...

- Y_x random variable parametrised by x
- $\mathbb{P}(Y_x)$ the probability of Y_x for given x
- Can only observe $\mathbb{P}(Y_x)$ via sampling
- $M(x) = \mathbb{E}(g(Y_x, x))$, for some $g(y, x)$
- **Problem**:
  Find θ such that $M(\theta) = 0$

Stochastic Approximation Algorithm

$$X_{n+1} = X_n + a_n g(y_n, x_n)$$

Robbins-Monro:

$$g(y, x) = \alpha - y$$

Law of large numbers:

$g(y, x) = y - x$ and Y independent of x

Banach fixed-point theorem:

$g(y, x) = y - x$ and $Y_x = \phi(x)$

Generalising...

- Y_x random variable parametrised by x
- $\mathbb{P}(Y_x)$ the probability of Y_x for given x
- Can only observe $\mathbb{P}(Y_x)$ via sampling
- $M(x) = \mathbb{E}(g(Y_x, x))$, for some $g(y, x)$
- **Problem**:
  Find θ such that $M(\theta) = 0$

Kolmogorov Strong Law of Large Numbers

Y a r. v. (independent of x)

Problem: Find x s.t. $\mathbb{E}(Y) = x$

Instance:

$g(y, x) = y - x$

$x_{n+1} = x_n + (y_n - x_n)/(n+1)$

(given samples $y_0, y_1, \ldots$)

SLLN: (x_n) converges to $\mathbb{E}(Y)$ a.s.

Generalising...

- Y_x random variable parametrised by x
- $\mathbb{P}(Y_x)$ the probability of Y_x for given x
- Can only observe $\mathbb{P}(Y_x)$ via sampling
- $M(x) = \mathbb{E}(g(Y_x, x))$, for some $g(y, x)$
- **Problem**:
  Find θ such that $M(\theta) = 0$

Banach Fixed-Point Theorem

Contraction mapping $\phi : \mathbb{R} \to \mathbb{R}$

Problem: Find x s.t. $\phi(x) = x$

Instance:

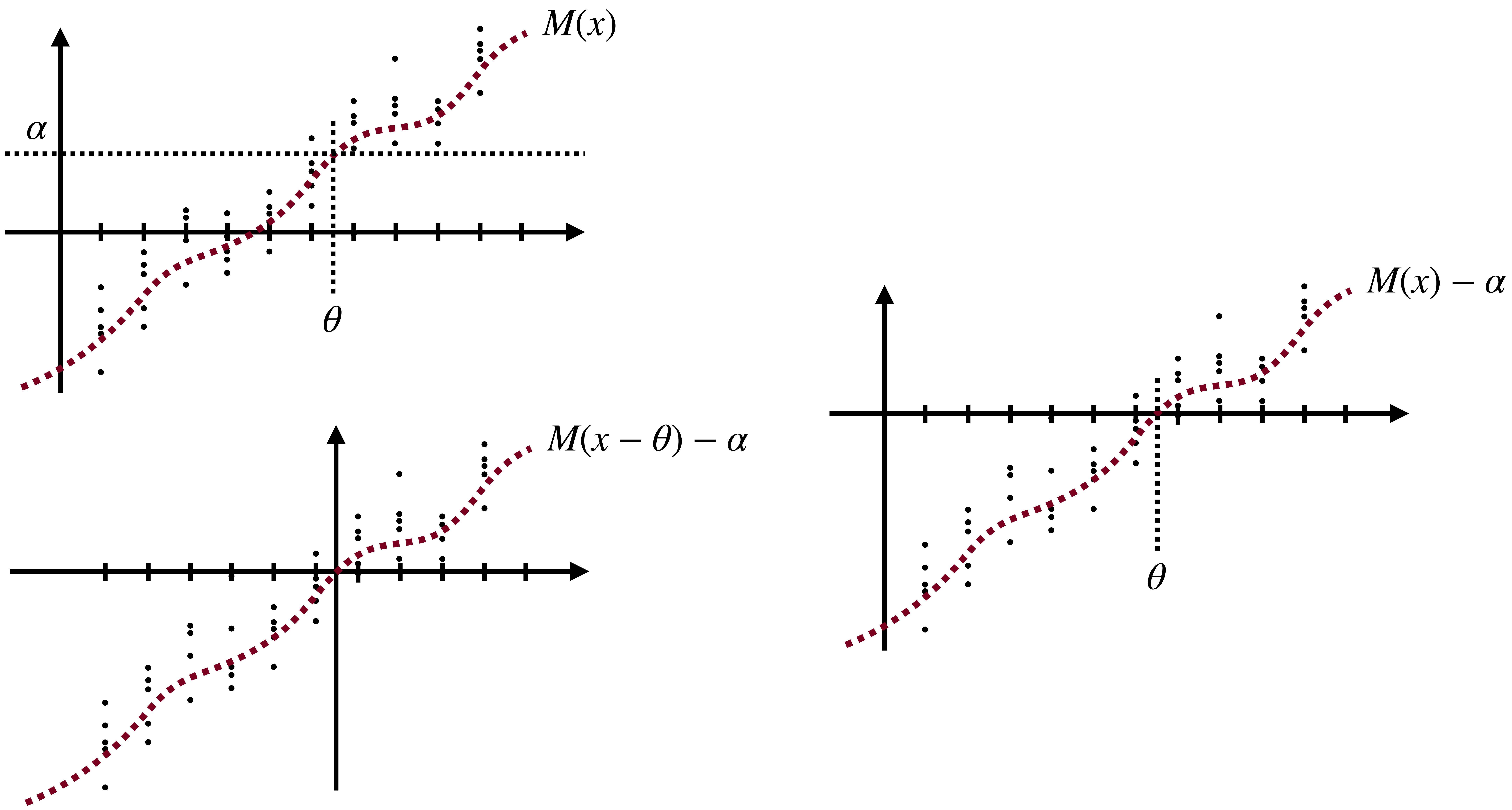
$Y_x = \phi(x)$

$g(y, x) = y - x$

$x_{n+1} = x_n + (\phi(x_n) - x_n)/(n + 1)$

BFT: (x_n) converges to f.p. of ϕ

Dvoretzky's Generalisation



$M(x)$
α
θ
$M(x-\theta)-\alpha$
$M(x)-\alpha$
θ

$$\boxed{x_{n+1} = x_n + a_n(\alpha - y_n)}$$

Robbins-Monro iteration

$$\boxed{x_{n+1} = x_n + a_n y_n}$$

w.l.g. we can assume $\alpha = 0$

$$\boxed{x_{n+1} = x_n + a_n y_n + a_n \mathbb{E}(Y_{x_n}) - a_n \mathbb{E}(Y_{x_n})}$$

add and remove $a_n \mathbb{E}(Y_{x_n})$

$$\boxed{x_{n+1} = x_n \underbrace{- a_n \mathbb{E}(Y_{x_n})}_{\text{"drift" deterministic}} + \underbrace{a_n(y_n - \mathbb{E}(Y_{x_n}))}_{\text{"noise" 0 on average}}}$$

rearrange

“drift”
deterministic

“noise”
0 on average

Dvoretzky (1956)

- Vast generalisation of Robbins-Monro and other results
- Iteration has two parts:
  - Deterministic part which is assume to converge to solution
  - Stochastic part with expected value of zero and bounded variance

Assumptions:

$T_n \colon \mathbb{R}^n \to \mathbb{R}$

$Y_1, Y_2, \ldots$ (dependent on history of $x_1, x_2, \ldots$)

- $\mathbb{E}[Y_n \mid X_1, \ldots, X_n] = 0$, w. p. 1
- $\Sigma\, \mathbb{E}[Y_n^2] < \infty$

There exists θ (can assume $\theta = 0$) such that

$$|T_n(\vec{x}) - \theta| \leq \max(\alpha_n, (1 + \beta_n)|x_n - \theta| - \gamma_n)$$

$\alpha_n, \beta_n, \gamma_n \in \mathbb{R}^+$ such that

$$\alpha_n \to 0, \;\; \Sigma\, \beta_n < \infty, \;\; \Sigma\, \gamma_n = \infty$$

Result:

The iteration

$$x_{n+1} = \underbrace{T_n(x_1, \ldots, x_n)}_{\textit{deterministic}} + \overbrace{Y_n}^{\textit{noise terms}}$$

converges a.s. to θ

THEOREM 1. (Dvoretzky). *Let* $\{X_n\}$, $\{T_n(X_1, \cdots, X_n)\}$, $\{Y_n(X_1, \cdots, X_n)\}$ *be sequences of real random variables with* X_1 *arbitrary and*

$$X_{n+1} = T_n(X_1, \cdots, X_n) + Y_n(X_1, \cdots, X_n). \tag{6}$$

Assume

$$E\{Y_n \mid X_1, \cdots, X_n\} = 0 \quad \text{w.p.1}, \tag{7}$$

$$\sum EY_n^2 < \infty, \tag{8}$$

and

$$|T_n| \leqq \max(\alpha_n, (1 + \beta_n)|X_n| - \gamma_n) \tag{9}$$

where α_n, β_n, γ_n *are positive numbers such that*

$$\alpha_n \to 0, \sum \beta_n < \infty, \qquad \sum \gamma_n = \infty. \tag{10}$$

Then $\boxed{X_n \to 0 \text{ w.p.1.}}$

Dvoretzky ⇒ Robbins-Monro

(Dvoretzky) $$x_{n+1} = T_n(x_1, \ldots, x_n) + Y_n(x_1, \ldots, x_n)$$

$$T_n(x_1, \ldots, x_n) = x_n + a_n \mathbb{E}(Z_{x_n})$$

$$Y_n(x_1, \ldots, x_n) = a_n(Z_{x_n} - \mathbb{E}(Z_{x_n}))$$

(Robbins-Monro) $$x_{n+1} = x_n + a_n Z_{x_n}$$

Derman-Sacks Proof (1959)

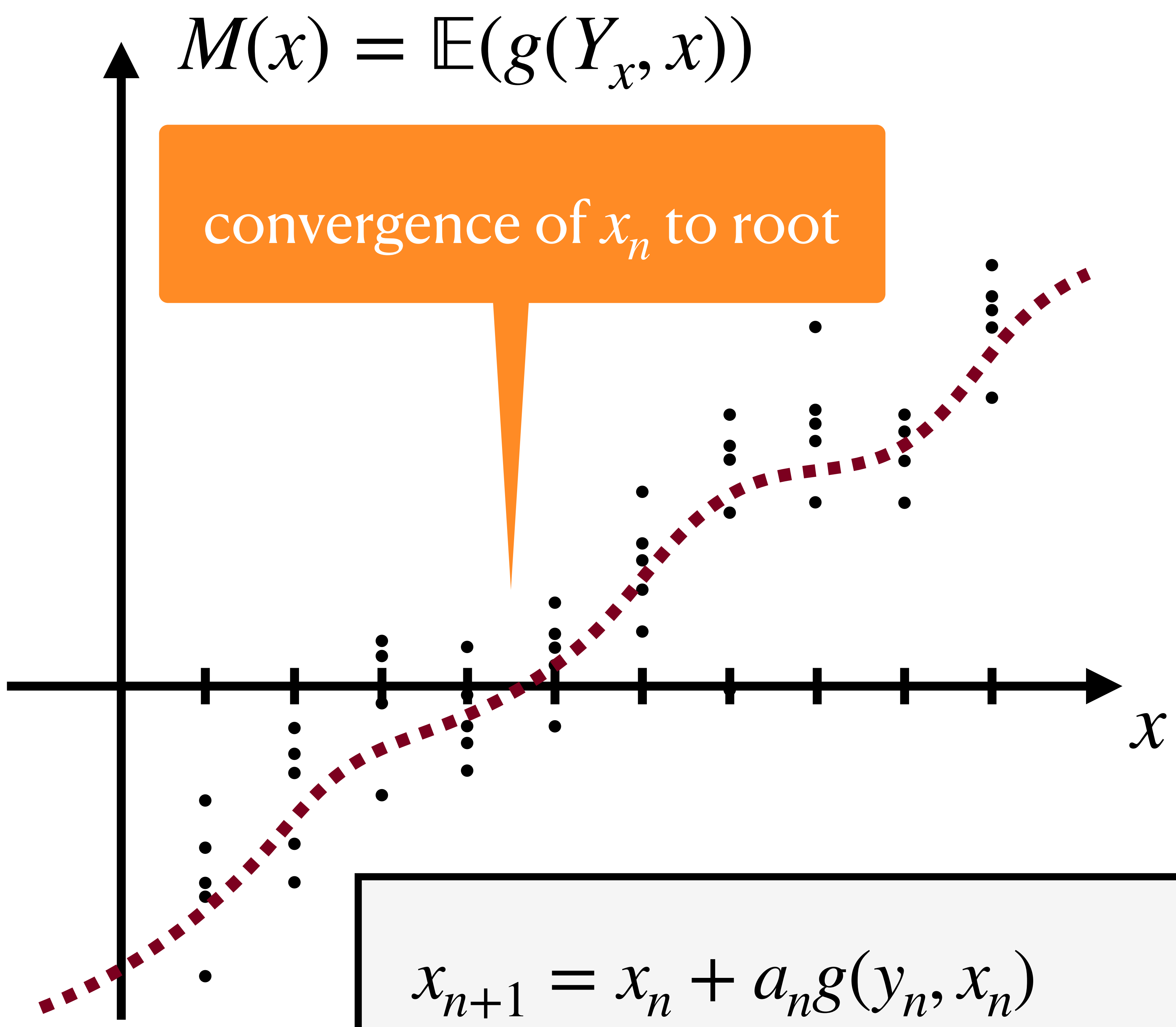
$^{*}X_n \text{ indep}.$

- Borel-Cantelli lemma (1st) $\Sigma\, \mathbb{P}(X_n) < \infty \Rightarrow \mathbb{P}(X_n \text{ i.o.}) = 0$
- Chebyshev inequality $\mathbb{P}(|X - \mu| \geq k\sigma) \leq 1/k^2$
- Abel's test $\Sigma\, a_n \text{ conv} \wedge b_n \text{ mon. and bounded} \Rightarrow \Sigma\, a_n b_n \text{ conv}$
- Slowdown lemma $\Sigma\, a_n \text{ conv} \Rightarrow \exists b_n (b_n \to 0 \wedge \Sigma\, a_n/b_n \text{ conv})$
- Kolmogorov inequality* $\mathbb{P}(\max_{1 \leq k \leq n} |X_1 + \ldots + X_k| \geq \lambda) \leq 1/\lambda^2 \mathrm{Var}(X_1 + \ldots + X_n)$
- Variance lemma* $\Sigma\, \mathbb{E}\, X_n^2 < \infty \Rightarrow \Sigma\, X_n \text{ a.s.}$
- "Lemma 1" about $\mathbb{R}$ converging and diverging sequences and series

Stochastic Approximation
(and Stochastic Gradient Descent)

Stochastic Approximation Methods

- Y_x random variable parametrised by x
- $\mathbb{P}(Y_x)$ the probability of Y_x for given x
- Can only observe $\mathbb{P}(Y_x)$ via sampling
- $g(y, x)$ a given function
- $M(x) = \mathbb{E}(g(Y_x, x))$
- **Problem**:
  Find x such that $M(\theta) = 0$



$x_{n+1} = x_n + a_n g(y_n, x_n)$

given samples $y_0, y_1, \ldots$ of $Y_{x_1}, Y_{x_2}, \ldots$

Stochastic Gradient Descent

loss function $L(y, x)$, x model param, y training data

Problem: Find x s.t. $\mathbb{E}(L(Y_x, x))$ is minimal

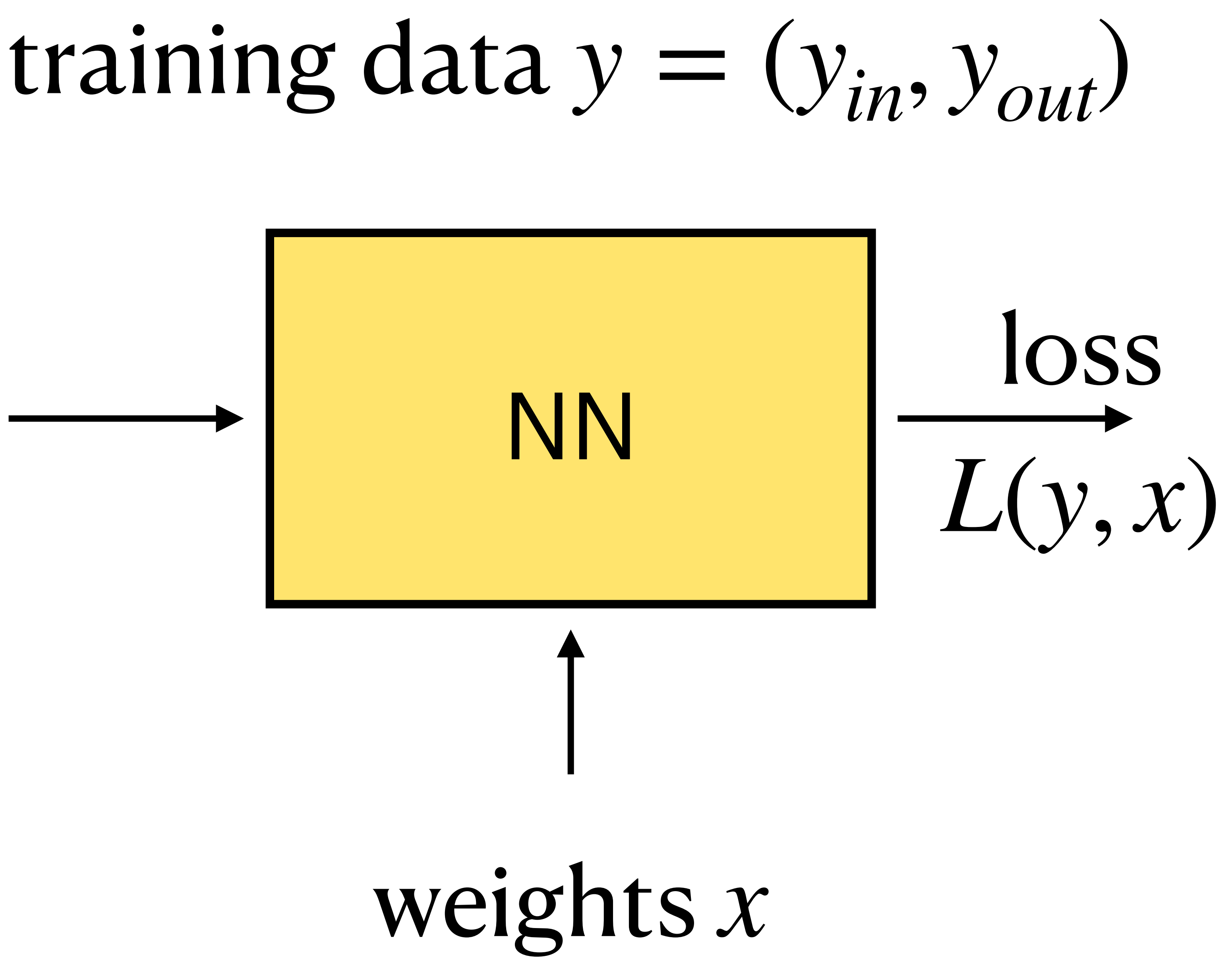
$g(y, x) = -\nabla_x L(y, x)$

Training set: $y_1, y_2, \ldots$ (samples of Y_x)

$x_{n+1} = x_n + a_n g(y_n, x_n)$

(a_n learning rate)

SGD: (x_n) converges to critical point of loss function



Quantitative Results

Qualitative vs Quantitative

- **Qualitative** (a.s.) convergence results say that with probability 1 the random variables X_n converge to some solution θ:

$$\mathbb{P}(\lim_{n\to\infty} X_n = \theta) = 1$$

- **Quantitative** (a.s.) convergence results produce a rate of a.s. convergence Φ: Given $\varepsilon, \lambda > 0$:

$$\mathbb{P}\left(\bigcap_{n\geq\Phi(\varepsilon,\lambda)} |X_n - \theta| < \varepsilon\right) > 1 - \lambda$$

Stochastic Approximation (1960 to 2026)

ODE Approach to SA (1970s)

- When we discretize (with step size a_n) an ordinary differential equation

$$(\dagger) \quad \dot{x}(t) = h(x(t))$$

  we get

$$x_{n+1} \simeq x_n + a_n h(x_n)$$

- SA behaves like a noisy, time-rescaled discretization of the ODE
- As $a_n \to 0$ the trajectory of x_n tracks the solution of $(\dagger)$

“Adaptive” Algorithms (1980s)

- Consider the iteration

$$x_{n+1} \simeq x_n + a_n h(x_n)$$

but focus on **convergence of the averages** (Polyak–Ruppert)

$$\overline{x_n} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Even with suboptimal step sizes a_n the sequence $\overline{x_n}$ achieves “**optimal asymptotic covariance convergence**”
- “**Tail averaging**” in practice (PyTorch and TensorFlow)

RL as SA with Markovian Noise (1990s)

- $s \in S$ are states, $a \in A$ are actions, $r \in \mathbb{R}$ are rewards, and policy π
- In reinforcement learning (RL) data comes from a **Markov process**:

$$s_n \to s_{n+1} \to s_{n+2} \to \dots$$

  and goal is to **maximise expected cumulative reward**

- Can only sample and approximate goal estimation
- Reinforcement learning algorithms are stochastic approximation with **Markovian noise** rather than i.i.d. noise

The End

Stochastic Approximation Methods

- **Iterated procedures** used to **approximate a target** value when target is unknown and observations are corrupted by **noise**
- Introduced by **Robbins-Monro** (1951)
- Robbins-Monro algorithms generalised by **Dvoretzky** (1956)
- Generalisation of (deterministic) approximation methods
- **Stochastic gradient descent** (in Machine Learning) is based on Stochastic Approximation Theory

$$x_{n+1} = x_n - \underbrace{\frac{f(x_n)}{f'(x_n)}}_{\text{correction term}}$$

... instead of $1/f'(x_n)$

approximation of $f(x_n)$

$$x_{n+1} = x_n - \underbrace{a_n y_n}_{\text{correction term}}$$