Thursday 11/17/05

Simulation & Learning

Recall: Last time, Prediction Error Method.

$$Y(n+1) = [H(\theta_0, z^{-1})U](n) + N(n+1)$$
$$\hat{Y}(n+1|n) = [H(\hat{\theta}(n), z^{-1})U](n), \quad \text{given estimate.}$$

Given $\hat{\theta} = \theta$ consider MSE. $\Gamma(\theta) = \frac{1}{2} E\left[\left(Y(n+1) - \hat{Y}(n+1|n)\right)^2\right]$.

Goal: Solve $\boxed{\nabla\Gamma(\theta) = -E\left[\left(Y(n+1) - \hat{Y}(n+1|n)\right)\nabla_\theta \hat{Y}(n+1|n)\right] = 0}$

Usual Gradient Algorithm

$$\hat{\theta}(k+1) = \hat{\theta}(k) - a\nabla\Gamma(\hat{\theta}(k)) \qquad \text{Not computable}$$

Stochastic Gradient Algorithm

$$\hat{\theta}(k+1) = \hat{\theta}(k) + a(k)\left[\left(Y(k+1) - \hat{Y}(k+1|k)\right)\phi(k)\right]$$

$$\phi(k) = \nabla_\theta \hat{Y}(k+1|k)\Big|_{\theta=\hat{\theta}(k)} = \left[\nabla_\theta H(\theta, z^{-1})U\right](k)\Big|_{\theta=\hat{\theta}(k)}.$$

Special case of stochastic Approximation

- Robbins + Monro
- Hirsch '89
- Benveniste et. al. '90
- Kushner + Yin '97

$\boxed{\text{Borkar + Meyn '00}}$

General set-up: $\theta^* \in \mathbb{R}^d$ unknown parameter

$f : \mathbb{R}^{d+m} \to \mathbb{R}^d$, $N \in \mathbb{R}^m$ random vector

$$E[f(\theta, N)] = 0 \quad \text{when } \theta = \theta^*.$$

Examples 1) Simulation $E[f(N) - \theta] = 0$, $\theta^* = \eta = \pi(f)$.

2) Optimization $f(\theta, N) = \nabla \Gamma(\theta, N)$

3) Fixed point equations Recall ACOE,

$$\min_u \left[c(x,u) + P_u h^*(x) \right] = h^*(x) + \eta^*, \quad x \in \mathsf{X}.$$

DCOE,
$$\min_u \left[c(x,u) + P_u h_\gamma^*(x) \right] = (1+\gamma) h_\gamma^*(x)$$

Define $Q(x,u) = (1+\gamma)^{-1} \left[c(x,u) + P_u h_\gamma^*(x) \right]$

$\therefore \quad \min_u Q(x,u) = h_\gamma^*(x)$

$\therefore \quad Q(x,u) = (1+\gamma)^{-1} \left[c(x,u) + \sum_y P_u(x,y) \left[\min_{u'} Q(y,u') \right] \right]$

(concave in Q)

Here $\Theta = \{ Q(x,u) : x \in \mathsf{X},\ u \in U(x) \}$, $N = \{ N(x,u) : \cdots \}$

$$f(\theta, N) = -Q(x,u) + (1+\gamma)^{-1} \left[c(x,u) + \min_{u'} Q(N, u') \right]$$

where $N(x,u) \sim X(k+1)$ when $X(k) = x$, $U(k) = u$.

$\longrightarrow$ SA $\equiv$ Q-learning

General S.A. algorithm

$$\Theta(k+1) = \Theta(k) + a_k f(\Theta(k), N(k+1)), \qquad k \geq 0.$$

Restrict to $\{N(k)\}$ iid.

Standard form: $g(\theta) = E[f(\theta, N)]$

$$\Theta(k+1) = \Theta(k) + a_k [g(\Theta(k)) + \Delta(k+1)]$$

$$\Delta(k+1) = f(\Theta(k), N(k+1)) - E[f(\Theta(k), N(k+1)) \mid \underbrace{\Theta_0^k, N_0^k}_{\mathcal{F}_k}]$$

Two canonical settings: Constant step-size $a_k \equiv a$

Tapering/Vanishing step-size

$$\boxed{\sum a_k = \infty, \quad \sum a_k^2 < \infty}$$

Assumptions

(A1) $g : \mathbb{R}^d \to \mathbb{R}^d$ is Lipschitz continuous,

$$\|g(x^1) - g(x^2)\| \leq b_g \|x^1 - x^2\| \qquad \text{all } x^1, x^2 \in \mathbb{R}^d.$$

(A2) $E[\|\Delta(n+1)\|^2 \mid \mathcal{F}_n] \leq \sigma_\Delta^2 (1 + \|\Theta(n)\|^2), \qquad n \geq 0.$

Issues. Stability & Convergence

Approach: ODE method

$$\dot{x} = g(x) \qquad \text{and other "fluid models".}$$

Simplest example is simulation: Monte-Carlo,

$$\theta(n+1) = \frac{1}{n+1} \sum_0^n f(N(k))$$

$$= \frac{1}{n+1} \left\{ n\,\theta(n) + f(N(n)) \right\}$$

$$= \theta(n) + \frac{1}{n+1} \left(f(N(n)) - \theta(n) \right)$$

More generally, take any sequence $\{a(n) \geq 0\}$ satisfying

$$\sum_0^\infty a_n = \infty, \quad \sum_0^\infty a_n^2 < \infty.$$

$$\theta(n+1) = \theta(n) + a_n \left[f(N(n)) - \theta(n) \right]$$

If $\{N\}$ is a nice Markov chain, or iid, then

$$\theta(n) \to \theta^* = \pi(f), \quad n \to \infty \quad \text{a.s.}$$

Fixed step-size algorithm: $\Theta(n+1) = \theta(n) + a\left[f(N(n)) - \Theta(n) \right]$

$$= (1-a)\,\theta(n) + a\, f(N(n))$$

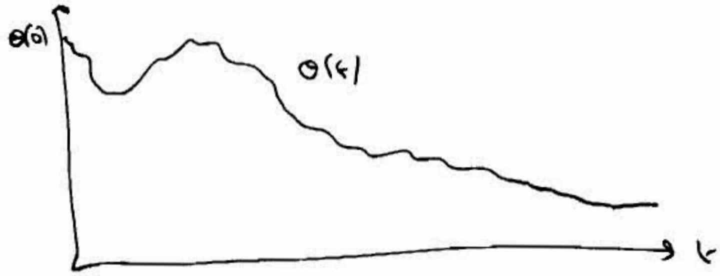
$$= (1-a)^{n+1}\theta(0) + \sum_{i=0}^n a(1-a)^{n-i} f(N(i)).$$

$$\boxed{E\left[\|\Theta(n) - \theta^*\|^2\right] \leq B_1(a) + B_2\left[1 + \|\theta(0)\|^2\right] e^{-\varepsilon(a) n}}$$

variance $O(a)$

This can be generalized to general algorithm.

Stability Considerations



Let $r \geq 1$, $\theta^r(k) = \frac{1}{r}\theta(k)$, with $\theta(0) = rx \in \mathbb{R}^d$.

$$\theta(k+1) = \theta(k) + a_k\left[g(\theta(k)) + \Delta(k+1)\right]$$

$$\theta^r(k+1) = \theta^r(k) + a_k\left[\tfrac{1}{r}g(\theta(k)) + \tfrac{1}{r}\Delta(k+1)\right]$$

Define $g_r(x) = \frac{1}{r}g(rx)$, $x \in \mathbb{R}^d$.

$$\theta^r(k+1) = \theta^r(k) + a_k\left[g_r(\theta^r(k)) + \tfrac{1}{r}\Delta(k+1)\right].$$

Motivates O.D.E. $\boxed{\dot{x} = g_r(x)}$ and $\boxed{\dot{x} = g_\infty(x)}$

provided $g_r(x) \to g_\infty(x)$ pointwise.

Proposition: Suppose the origin is locally asymptotically stable for $\dot{x} = g_\infty(x)$. Then it is globally exponentially asymptotically stable

Proof: g_∞ is homogeneous,

$$g_\infty(tx) = \lim_{r\to\infty} \frac{1}{r} g(rtx) = t\, g_\infty(x).$$

Let $\varepsilon > 0$ s.t. $x(t) \to 0$ uniformly for $x(0) \in \overline{B(\varepsilon)} = \{x : \|x\| \leq \varepsilon\}$.

Thus, we can find $T > 0$ s.t. $\|x(T)\| \leq \varepsilon/2$ whenever $\|x(0)\| \leq \varepsilon$.

Consider any solution, and write $\bar{x}(t) = \varepsilon \dfrac{x(t)}{\|x(0)\|}$.

$\Rightarrow \|\bar{x}(T)\| \leq \varepsilon/2 \quad \Rightarrow \quad \|x(T)\| \leq \frac{1}{2}\|x(0)\|$

[Note: can also project]

Constant step-size algorithm : Main ideas from Borkar + Meyn

$\theta(t)$ is a Markov chain.

If g_∞ defines stable vector field then for large r, small a,

$$E[\|\theta^r(T)\|^2] \leq \tfrac{3}{4}\|\theta^r(0)\|^2 \quad \text{when } \|\theta^r(0)\| \geq 1.$$

Proved by relating stability of g_∞ to stability of g_r ...

$$\Rightarrow \quad E[\|\theta(T)\|^2] \leq \tfrac{3}{4}\|\theta(0)\|^2 \quad \text{when } \|\theta(0)\| \geq r.$$

Lyapunov drift condition. Provided a density condition holds,

Proposition There exists $a_0 > 0$ s.t. for all $0 < a \leq a_0$

$$E[\|\theta(n) - \theta^*\|^2] \leq B_1(a) + B_2[1 + \|x\|^2]\, e^{-\varepsilon_0(a) n}.$$

where $B_1(a) \to 0$, $\varepsilon_0(a) \to 0$, $a \to 0$.

$B_1(a) = E_\pi[\|\theta(0) - \theta^*\|^2]$.

"Mean - Variance trade-off"

Vanishing step-size algorithm

Preliminaries: $\{\Delta(k) : k \geq 1\}$ is a martingale - difference sequence,

$$E[\Delta(k+1) \mid \mathcal{F}_k] = 0, \qquad E[\|\Delta(k+1)\|^2 \mid \mathcal{F}_k] \leq \sigma_\Delta^2 (1 + \|\theta(k)\|^2)$$

(A2)

$$(*) \quad \theta(k+1) = \theta(0) + \sum_{i=0}^{k} a_i g(\theta(i)) + M(k+1),$$

where $M(\cdot)$ is a martingale

$$M(k+1) = \sum_{i=0}^{k} a_i \Delta(i+1), \qquad E[M(k+1) \mid \mathcal{F}_k] = M(k).$$

$$E[\|M(k+1)\|^2] = \sum_{i=0}^{k} a_i^2 E[\|\Delta(i+1)\|^2]$$ bounded, if $E[\|\Delta(i)\|^2]$ is bounded.

Under (A2): require $E[\|\theta(k)\|^2]$ bounded.

Boundedness is established under stability of O.D.E

$$\dot{x} = g_\infty(x).$$

Approach to convergence: $(*)$ looks like a discrete approximation to the solution to the ODE

$$\dot{x} = g(x)$$

which we assume has unique a. stable equilibrium θ^*.

Also, Assume: $\theta(k) \in H$ compact for all $k \geq 0$.

<u>Time scale</u>: $t(n) = \sum_{i=0}^{n-1} a(i) \to \infty$, as $n \to \infty$.

Fix $T > 0$ and define $T(0) = 0$,

$$T(n+1) = \min\big(t(j) : t(j) > T(n) + T\big), \quad n \geq 0.$$

We have $T(n+1) - T(n) \geq T$ for each n, and

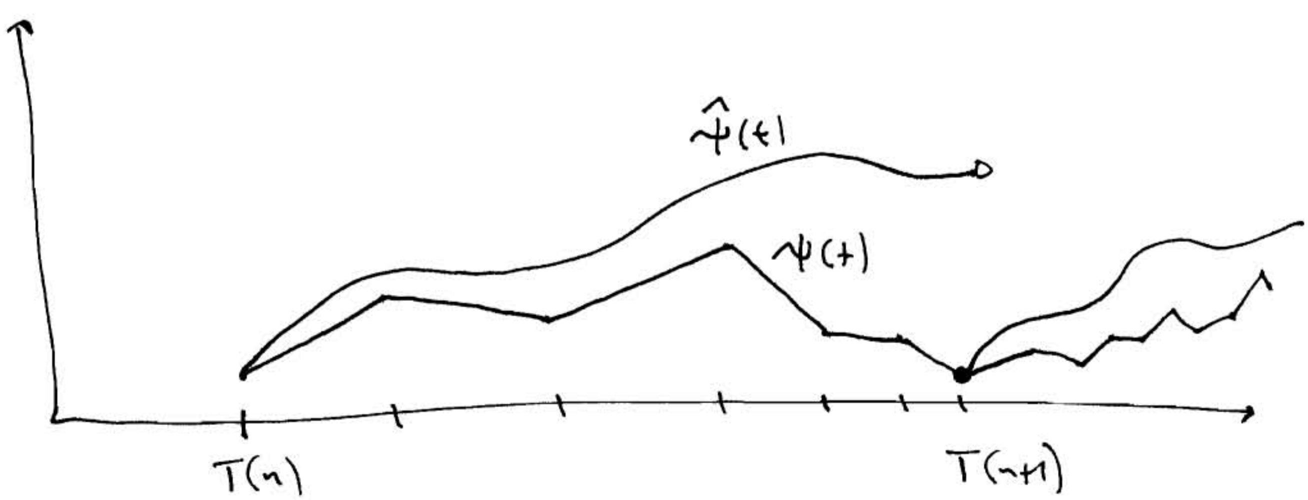
$$T(n+1) - T(n) \to T, \quad n \to \infty.$$

<u>Two processes to be compared</u>: Both defined for $t \in \mathbb{R}_+$,

$\psi(t)$: piecewise linear, $\psi(t(n)) = \theta(n)$.

$\hat{\psi}(t)$: piecewise continuous, on each interval $[T(n), T(n+1))$, it is the solution to the ODE

$$\frac{d}{dt}\hat{\psi}(t) = g(\hat{\psi}(t)); \quad \hat{\psi}(T(n)) = \psi(T(n))$$



$\hat{\psi}$ takes "jump" at times $\{T(n)\}$.

For comparison:

(i) Fix $\varepsilon > 0$, and let $B(\varepsilon)$ denote open ball of radius ε, centered at θ^*

(ii) Find $0 < \delta < \varepsilon$ such that $x(t) \in B(\varepsilon)$ for $t \geq 0$ when $x(0) \in B(\delta)$ (for ODE!)

(iii) Find $T > 0$ so large that $x(t) \in B(\delta/2)$ for $t \geq T$ when $x(0) \in H$.

<u>Note</u>: Under (iii) we have $\hat{\varphi}(T(n)-) \in B(\delta/2)$ for each $n \geq 1$.

Next: we bound $\|\varphi(t) - \hat{\varphi}(t)\|$ for $t \geq 0$.

This uses Lipschitz continuity of g, and

<u>Bellman - Gronwall Lemma</u> Suppose $\{A(t) : 0 \leq t \leq T\}$ is non-negative, and satisfies

$$A(t) \leq A(0) + b\int_0^t A(s)\, ds, \qquad 0 \leq s \leq T.$$

Then,

$$A(t) \leq A(0)e^{bt}, \qquad 0 \leq s \leq T.$$

We have for each $t(k) > T(n)$,

$$\psi(t(k)) = \psi(T(n)) + \sum_j a_j g(\psi(t(j))) + M(t(k)) - M(T(n))$$

$\{j : T(n) \le t(j) < t(k)\}$.

With some work,

$$\psi(t) = \psi(T(n)) + \int_{T(n)}^{t} g(\psi(t))\,dt + \mathcal{E}(T(n), t)$$

such that $\lim_{n\to\infty}\left(\sup_{T(n)\le t\le T(n+1)} \|\mathcal{E}(T(n), t)\|\right) = 0.$

Also, by definition,

$$\hat{\psi}(t) = \psi(T(n)) + \int_{T(n)}^{t} g(\hat{\psi}(t))\,dt, \quad T(n) \le t < T(n+1).$$

So, with b_0 equal to the Lipschitz constant,

$$\|\psi(t) - \hat{\psi}(t)\| \le b_0 \int_{T(n)}^{t} \|\psi(s) - \hat{\psi}(s)\|\,ds + e(n)$$

$$T(n) \le t < T(n+1),$$

where $e(n) = \sup_{T(n)\le t< T(n+1)} \|\mathcal{E}(T(n), t)\|$.

To place this in the form of the B.G. Lemma define,

$$A(t) = \max\left(\|\psi(t) - \hat{\psi}(t)\|, e(n)\right),$$

$$T(n) \le t < T(n+1).$$

$$A(t) \le \max\left\{ b_0 \int_{T(n)}^{t} \| \psi(s) - \hat{\psi}(s) \| \, ds + e(n), \; e(n) \right\}$$

$$\le b_0 \int_{T(n)}^{t} A(s) \, ds + \underset{\equiv A(0)}{\underset{\uparrow}{e(n)}}$$

$$\therefore \quad \max\left(\| \psi(t) - \hat{\psi}(t) \|, e(n) \right) \le e(n) \, e^{b_0 (t - T(n))}$$

$$T(n) \le t < T(n+1)$$

So, $\sup\limits_{T(n) \le t < T(n+1)} \| \psi(t) - \hat{\psi}(t) \| \longrightarrow 0, \quad n \to \infty.$

In particular, $\psi(T(n)) \in B(\delta)$ for all large n.

$\implies \hat{\psi}(t) \in B(\varepsilon)$ for all large t

$$\implies \limsup_{t \to \infty} \| \psi(t) - \theta^* \| \le \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows that $\theta(n) \to \theta^*$ provided $\{\theta(n)\}$ is bounded.