Adaptive maximization of social welfare

Dennis Hein   Maximilian Kasy

Department of Economics, University of Oxford

Overview

- Problem: Repeatedly choose policy parameters to **maximize social welfare**, the weighted sum of utility.
- Vs. multi-armed bandits: **Utility is not observed**, but needs to be indirectly inferred as equivalent variation.
- Vs. standard optimal tax theory: **Response functions need to be learned** through policy choices.
- Proposed algorithm: Combine optimal tax theory, Gaussian process priors, random Fourier features, and Thompson sampling.

Optimal taxation

- Individuals t arrive sequentially.
- They choose $Y_t \in \mathbb{R}$ subject to a linear tax rate X_t.
- Taxes owed: $Y_t \cdot X_t$.
- Response function: $Y_t = g(X_t, U_t)$.
- Average response function $m(x) = E[g(x, U_t)] = E[Y_t | X_t = x]$.

Social welfare

- Expected tax revenues: $m(x) \cdot x$.
- Private welfare: $-\int_0^x m(x')dx'$. (Envelope theorem $\Rightarrow$ consumer surplus!)
- Welfare weight $\lambda \Rightarrow$ social welfare

$$s(x) = m(x) \cdot x - \lambda \int_0^x m(x')dx'. \tag{1}$$

Gaussian process prior and posterior

- Gaussian process prior:

$$m(\cdot) \sim GP(\mu(\cdot), C(\cdot, \cdot)). \tag{2}$$

- Posterior of social welfare:

$$E[s(x)|Y_t, X_t] = \nu(x) + D_t(x) \cdot \left[C_t + \sigma^2 I\right]^{-1} \cdot (Y_t - \mu_t),$$
$$\text{Var}(s(x)|Y_t, X_t) = \text{Var}(s(x)) - D_t(x) \cdot \left[C_t + \sigma^2 I\right]^{-1} \cdot D_t^\top(x),$$
$$\nu(x) = E[s(x)] = x \cdot \mu(x) - \lambda \int_0^x \mu(x')dx',$$
$$D(x, x') = \text{Cov}(s(x), m(x'))) = x \cdot C(x, x') - \lambda \cdot \int_0^t C(x, x')dx.$$

Algorithm: Thompson sampling for social welfare

Require: The history of tax rates and individual responses, X_{t-1}, Y_{t-1}. Hyper-parameters ρ, τ^2, σ^2.

Sample $j = 1, \ldots, k$ i.i.d. draws $\theta_{j1} \sim N(0, \rho)$ and $\theta_{j0} \sim U[0, 2\pi]$.

Calculate the matrix Φ_{t-1} with entries $\sqrt{\frac{2\tau^2}{k}} \cos(x_{t'} \cdot \theta_{j1} + \theta_{j0})$.

Sample one draw of the vector $\hat{\omega}_t$ from the distribution

$$N\left(\left(\Phi_{t-1}^T \Phi_{t-1} + \sigma^2 I\right)^{-1} \cdot \Phi_{t-1}^T Y_{t-1}, \quad \left(\Phi_{t-1}^T \Phi_{t-1} + \sigma^2 I\right)^{-1} \cdot \sigma^2\right).$$

Set a starting value $x = X_{t-1}$.

while Convergence criterion for Newton's method is not achieved **do**

  Evaluate $\hat{s}'_t(x)$ and $\hat{s}''_t(x)$ for $\hat{s}_t(x) = \sum_{j=1}^k \hat{\omega}_{tj} \cdot \left[\sqrt{\frac{\tau^2}{k}} \psi_j(x)\right]$, where

$$\psi'_j(x) = \phi'_j(x) \cdot x + (1 - \lambda) \cdot \phi_j(x), \qquad \phi_j(x) = \sqrt{2} \cos(x \cdot \theta_{j1} + \theta_{j0})$$
$$\psi''_j(x) = \phi''_j(x) \cdot x + (2 - \lambda) \cdot \phi'_j(x), \qquad \phi'_j(x) = -\sqrt{2}\theta_{j1} \sin(x \cdot \theta_{j1} + \theta_{j0})$$
$$\phi''_j(x) = -\sqrt{2}\theta_{j1}^2 \cos(x \cdot \theta_{j1} + \theta_{j0}).$$

  Update $x \leftarrow x - \frac{\hat{s}'_t(x)}{\hat{s}''_t(x)}$.

end while

return $X_t = x$.

Algorithm explained

1) Thompson sampling

- Sampling distribution of X_t := posterior distribution of $x^* = \text{argmax}_x\, s(x)$.
- Implementation: Sample $\hat{s}_t(\cdot)$ from the posterior for $s(\cdot)$.
- Set $X_t = \text{argmax}_x\, \hat{s}_t(x)$.

2) Random Fourier features

- Sampling a function and maximizing it is numerically challenging.
- We can approximate by a ridge regression: For ω_j i.i.d. $N(0, 1)$,

$$m(x) \approx \sum_{j=1}^k \omega_j \cdot \left[\sqrt{\frac{\tau^2}{k}} \phi_j(x)\right]. \tag{3}$$

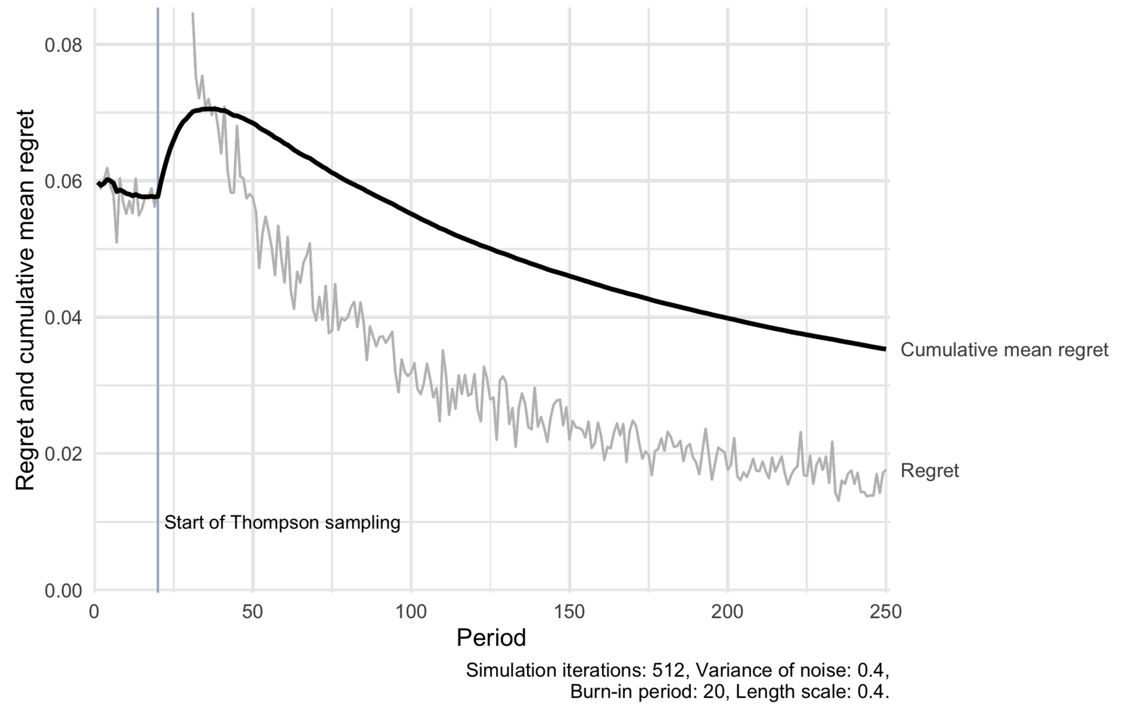
- Implied social welfare:

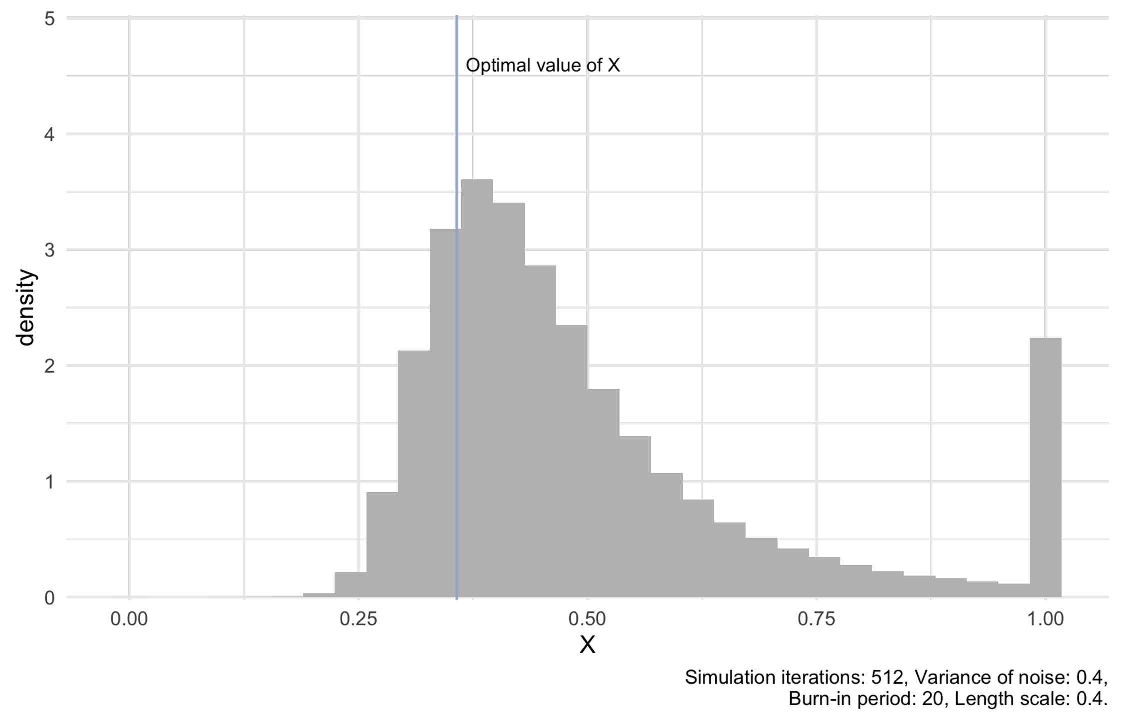
$$s(x) = \sum_{j=1}^k \omega_j \cdot \left[\sqrt{\frac{\tau^2}{k}} \psi_j(x)\right], \qquad \psi_j(x) = \phi_j(x) \cdot x - \lambda \int_0^x \phi_j(x')dx'.$$

$\Rightarrow$ Only need to obtain one draw of the ω_j from the posterior, and hold it constant during optimization of $\hat{s}_t(x)$.

- How to find $\phi_j(x)$? By Fourier transform of the squared-exponential kernel, $\phi_j(x) = \sqrt{2} \cos(x \cdot \theta_{j1} + \theta_{j0})$, with $\theta_{j1} \sim N(0, \rho)$ and $\theta_{j0} \sim U[0, 2\pi]$.

Simulations





Next steps (1): Basic income experiment

- With the NGO "Mein Grundeinkommen" in Germany.
- Participants will be assigned to different levels of transfer size and marginal tax rate (3×3 combinations).
- Assignment shares will be updated in waves.
- A parametric model of responses might be used for Thompson.

Next steps (2): Lower and upper regret bounds

- This setting has some relationship to adaptive choice of reserve prices in auctions, and to bilateral trade.
- Lower regret bounds for any algorithm, and upper bounds for specific algorithms, will be derived for the stochastic and adversarial settings.