Monte Carlo

Sinan Yıldırım

Giriş
Örneklerin ortalaması
Monte Carlo

Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi

Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings

Önem örneklemesi

Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Monte Carlo Yöntemleri

Sinan Yıldırım

MDBF, Sabancı Üniversitesi

August 12, 2018

İçindekiler

1. Giriş
   - Örneklerin ortalaması
   - Monte Carlo
2. Kesin örnekleme yöntemleri
   - Tersini alma yöntemi
   - Dönüştürme yöntemi
   - Birleştirme yöntemi
   - Reddetme örneklemesi
3. Markov zinciri Monte Carlo
   - Gibbs örneklemesi
   - Metropolis-Hastings
4. Önem örneklemesi
5. Sıralı Monte Carlo
   - Sıralı önem örneklemesi
   - Parçacık süzgeci

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Giriş

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Örneklerin ortalaması

Örneklerin Ortalaması

Bir $\mathcal{X} \subset \mathbb{R}^{d_x}$, $d_x \geq 1$ kümesinden $N \geq 1$ tane *rassal örnek* verilmiş olsun:

$$X^{(1)}, \ldots, X^{(N)}.$$

Örneklerin *bağımsız ve özdeş dağılımlı* olduğunu ve π olasılık dağılımından geldiğini varsayalım:

$$X^{(1)}, \ldots, X^{(N)} \overset{\text{i.i.d.}}{\sim} \pi.$$

Ayrıca π dağılımı bilinmiyor olsun.

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

π'ye göre ortalama değer

X'in π dağılımına göre beklentisini $X^{(1)}, \ldots, X^{(N)}$ örneklerini kullanarak yaklaşık olarak nasıl hesaplayabiliriz?

$\pi(x)$ olasılık yoğunluk fonksiyonu ise:

$$\mathbb{E}_\pi(X) = \int_{\mathcal{X}} x\pi(x)dx.$$

$\pi(x)$ olasılık kütle fonksiyonu ve X, $x_1, x_2, \ldots$ değerlerini alıyorsa:

$$\mathbb{E}_\pi(X) = \sum_i x_i\pi(x_i).$$

Makul bir kestirim:

$$\mathbb{E}_\pi(X) \approx \frac{1}{N}\sum_{i=1}^{N} X^{(i)}.$$

Genel fonksiyonların beklenti değeri

Şimdi de $\varphi : \mathcal{X} \to \mathbb{R}$ fonksiyonunun π'ye göre beklentisini inceleyelim.

$$\pi(\varphi) := \mathbb{E}_{\pi}(\varphi(X)) = \int_{\mathcal{X}} \varphi(x)\pi(x)dx.$$

Bu beklentinin kestirimi:

$$\mathbb{E}_{\pi}(\varphi(X)) \approx \frac{1}{N} \sum_{i=1}^{N} \varphi(X^{(i)}).$$

Örnek: $\varphi(x) = x^2$, $\varphi(x) = \log x$, vs.

$\varphi(X) = X$ bizi ilk probleme geri götürür.

Bir kümenin olasılığı

$A \subseteq \mathcal{X}$ şeklinde bir küme verilmiş olsun.

$$\pi(A) := \mathbb{P}(X \in A)$$

İşaret fonksiyonu: $\mathbb{I}_A : \mathcal{X} \to \{0, 1\}$

$$\mathbb{I}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Üstteki olasılık $\varphi = \mathbb{I}_A$ fonksiyonunun beklenti değeri olur:

$$\begin{aligned} \mathbb{E}_\pi(\mathbb{I}_A(X)) &= \int_{\mathcal{X}} \mathbb{I}_A(x)\pi(x)dx \\ &= \int_A \pi(x)dx \\ &= \mathbb{P}(X \in A). \end{aligned}$$

Bu olasılığa örnekler kullanılarak

$$\mathbb{P}(X \in A) \approx \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_A(X^{(i)}).$$

şeklinde yaklaşılabilir.

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Monte Carlo

Monte Carlo: Ana fikir

Şimdi şu senaryoyu düşünelim: π dağılımını biliyoruz, ama π'den gelen örneklerimiz yok.

- π'den istediğimiz kadar bağımsız örnek üretebiliyoruz.
- $\pi(\varphi)$'yi hesaplayamıyoruz.

Bu durumda $\pi(\varphi)$'ye nasıl yaklaşılabilir?

Eğer $X^{(1)}, \dots, X^{(N)} \overset{\text{i.i.d.}}{\sim} \pi$ örneklerini kendimiz üretirsek, ilk probleme geri döneriz.

Bu basit fikir, Monte Carlo yöntemlerinin ana fikridir.

Monte Carlo'nun gerekçelendirilmesi - yansızlık

$\pi(\varphi)$'nin Monte Carlo kestirimini $\pi_{\text{MC}}^N(\varphi)$ ile gösterelim:

$$\pi_{\text{MC}}^N(\varphi) = \frac{1}{N}\sum_{i=1}^{N}\varphi(X^{(i)}).$$

Herhangi bir $N \geq 1$ için, $\pi_{\text{MC}}^N(\varphi)$'nın beklenti değeri:

$$\begin{aligned}
\mathbb{E}\left(\pi_{\text{MC}}^N(\varphi)\right) &= \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\varphi(X^{(i)})\right). \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}_\pi(\varphi(X^{(i)})) \\
&= \frac{1}{N}N\mathbb{E}_\pi(\varphi(X)) \\
&= \mathbb{E}_\pi(\varphi(X)) = \pi(\varphi).
\end{aligned}$$

Ancak, yansızlık tek başına yeterli bir özellik değildir.

Monte Carlo'nun gerekçelendirilmesi - Büyük sayılar kanunu

$$\pi_{\mathrm{MC}}^{N}(\varphi) = \frac{1}{N}\sum_{i=1}^{N}\varphi(X^{(i)}).$$

Büyük sayılar kanunu: Eğer $|\pi(\varphi)| < \infty$ ise $\pi_{\mathrm{MC}}^{N}(\varphi)$'nin $\pi(\varphi)$'ye yakınsar:

$$\pi_{\mathrm{MC}}^{N}(\varphi) \xrightarrow{a.s.} \pi(\varphi), \quad \text{as } N \to \infty.$$

Monte Carlo'nun gerekçelendirilmesi - Merkezi limit teoremi

$\pi_{\mathrm{MC}}^{N}(\varphi)$'nin varyansı:

$$\mathbb{V}\left[\pi_{\mathrm{MC}}^{N}(\varphi)\right] = \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{V}_{\pi}\left[\varphi(X^{(i)})\right] = \frac{1}{N}\mathbb{V}_{\pi}\left[\varphi(X)\right].$$

Buradan, $\mathbb{V}_{\pi}\left[\varphi(X)\right]$ sonlu olduğu sürece $\pi_{\mathrm{MC}}^{N}(\varphi)$'nin doğruluğunun N ile arttığı söylenebilir.

Merkezi limit teoremi: Eğer $\mathbb{V}_{\pi}\left[\varphi(X)\right] < \infty$ ise

$$\sqrt{N}\left[\pi_{MC}^{N}(\varphi) - \pi(\varphi)\right] \xrightarrow{d} \mathcal{N}\left(0, \mathbb{V}_{\pi}\left[\varphi(X)\right]\right) \quad \text{as } N \to \infty.$$

Monte Carlo gerekçelendirilmesi - Deterministik integraller

$\pi(\varphi)$'nin hesaplanması için bir takım belirlemeci integral teknikleri de vardır;

Ancak bu teknikler X'in boyutu d_x büyüdükçe kötüleşir.

Monte Carlo yaklaşımının başarımı d_x'ten bağımsızdır.

$$\mathbb{V}\left[\pi^{N}_{\mathrm{MC}}(\varphi)\right] = \frac{1}{N}\mathbb{V}_{\pi}\left[\varphi(X)\right].$$

İleri yöntemlere ihtiyaç

Çoğu problemde, tek sorun integralin alınamazlığı değil.

- π'den örnekleme yapmak homojen dağılım kadar kolay olmayabilir.

  Bunun için bir takım kesin örnekleme yöntemleri geliştirilmiştir. Örnek: tersini alma yöntemi, reddetme örneklemesi, kompozisyon, vs

- π'den örnekleme yapmak imkansız olabilir.

  Örnek: Bayesçi çıkarımdaki sonsal dağılım: X bilinmeyen değişkenininin $Y = y$ verisi verildiğindeki sonsal dağılımı

$$\pi(x) := p_{X|Y}(x|y) = \frac{p_X(x)p_{Y|X}(y|x)}{\int p_X(x')p_{Y|X}(y|x')dx'} = \frac{p_{X,Y}(x,y)}{\int p_{X,Y}(x',y)dx'}$$
$$\propto p_X(x)p_{Y|X}(y|x)$$

  Bu tür dağılımlardan *yaklaşık* örnekler elde etmek için yazında bir çok yöntem var, örn: Markov zinciri Monte Carlo, Sıralı Monte Carlo, vs.

Giriş
Örneklerin ortalaması
Monte Carlo

Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi

Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings

Önem örneklemesi

Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Kesin örnekleme yöntemleri

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Sözde-rassal sayı

Çıkış noktası olarak, bilgisayarımızın homojen dağılımdan örnekler üretebildiğini varsayacağız.

$$U \sim \text{Unif}(0, 1)$$

Bu üretilen sayılar belirlenimci yöntemlerle üretilir; bu sebeple bu sayılara sözde-rassal sayı denir.

Soru: Elimizde Unif(0, 1) dağılımından gelen sayılar olsun. Bu sayıları kullanarak herhangi bir π dağılımından nasıl örnek üretebiliriz?

Tersini alma yöntemi

Tersini alma yöntemi

$X \sim \pi$ rassal değişkeninin kümülatif dağılım fonksiyonu:

$$F(x) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R}.$$

F'nin genelleştirilmiş tersi:

$$G(u) := \inf\{x \in \mathcal{X} : F(x) \geq u\}.$$

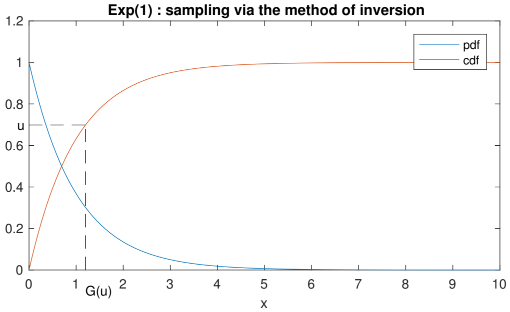
Homojen dağılmış sayılar ve G kullanılarak $X \sim \pi$ elde edilebilir.

$$U \sim \text{Unif}(0, 1) \Rightarrow G(U) \sim \pi$$

Örnek: Üssel dağılım

$X \sim \text{Exp}(\lambda)$, $\lambda > 0$, olasılık yoğunluk dağılımı

$$\pi(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{else} \end{cases}, \quad u = F(x) = \begin{cases} \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{else} \end{cases}.$$

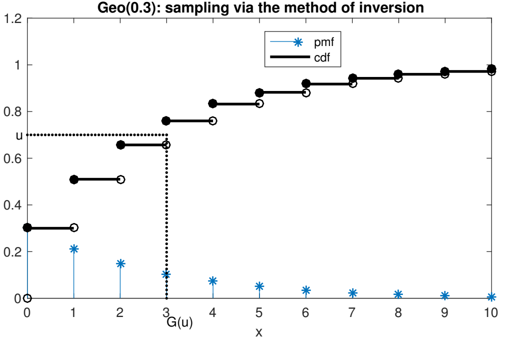


O halde, $\text{Exp}(\lambda)$'dan $U \sim \text{Unif}(0,1)$ ve $X = -\log(1-U)/\lambda$ şeklinde örnek üretebiliriz.

Örnek: Geometrik dağılım

$X \sim \text{Geo}(\rho)$, olasılık kütle fonksiyonu

$$\pi(x) = (1-\rho)^x \rho, \quad x = 0, 1, 2 \dots, \quad F(x) = 1 - (1-\rho)^{x+1}.$$



O halde, $\text{Geo}(\rho)$'dan $U \sim \text{Unif}(0,1)$ ve $X = \left\lceil \frac{\log(1-U)}{\log(1-\rho)} - 1 \right\rceil$ şeklinde örnek üretilebilir.

Dönüştürme yöntemi

Dönüştürme yöntemi: basit durum

Tersini alma yöntemi U'dan $X = G(U)$'ya bir çeşit dönüştürme olarak görülebilir.

Daha genel olarak, uygun bir g fonksiyonuyla bir dağılımdan diğerine geçilebilir.

Basit örnek: $X \sim \text{Unif}(a, b)$ üretmek için, $U \sim \text{Unif}(0, 1)$ üretip U'yu

$$X = g(U) := (b - a)U + a.$$

şeklinde dönüştürebiliriz.

Uygulama: $\mathcal{N}(0,1)$ için Box-Muller yöntemi

Standart Gauss dağılımı (normal dağılım) $\mathcal{N}(0,1)$'ın olasılık yoğunluk fonksiyonu

$$\phi(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Kümülatif dağılım fonksiyonunun tersini almak kolay değil. Alternatif olarak, dönüştürme kullanacağız.

1. İlk önce

$$R \sim \mathrm{Exp}(1/2), \quad \Theta \sim \mathrm{Unif}(0, 2\pi).$$

üretilir.

2. Sonra

$$X_1 = \sqrt{R}\cos(\Theta), \quad X_2 = \sqrt{R}\sin(\Theta)$$

dönüşümü ile $X_1, X_2 \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ elde edilir.

Çokdeğişkenli Gauss dağılım

$n \times 1$ boyutlu çok değişkenli Gauss dağılımını $X \sim \mathcal{N}(\mu, \Sigma)$ şeklinde gösterelim.

$\mu = \mathbb{E}(X)$, $n \times 1$ *ortalama vektörüdür*.

$$\Sigma = \text{Cov}(X) = \mathbb{E}[(X - \mu)(X - \mu)^T]$$

ise $n \times n$ simetrik ve kesin artı *kovaryansa* matrisidir. Bu matrisin (i, j)'inci elemanı

$$\sigma_{ij} = \text{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathbb{E}(X_i X_j) - \mu_i \mu_j$$

Olasılık yoğunluk fonksiyonu

$$\phi(x; \mu, \Sigma) = \frac{1}{|2\pi\Sigma|} \exp\left\{-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right\}$$

Burada, $|\cdot|$ determinantı simgeler.

Çokdeğişkenli Gauss dağılımı: örnekleme

$n \times 1$ boyutlu μ vektörü ve $n \times n$ kesin artı Σ matrisi verildiğinde, $X \sim \mathcal{N}(\mu, \Sigma)$ nasıl üretilir?

1. Önce $R_1, \ldots, R_n \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ üretilir böylece
$$R = (R_1, \ldots, R_n) \sim \mathcal{N}(0_n, I_n)$$
sağlanmış olur.
2. Sonra, Cholesky ayrıştırması kullanılarak
$$\Sigma = AA^T$$
eşitliğini sağlayan A matrisi bulunur.
3. Son olarak
$$X = AR + \mu$$
değişkeni üretilir.

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Birleştirme yöntemi

Birleştirme yöntemi: Sıradüzenli modeller

$\mathcal{Z}$ kümesinden değer alan ve $Z \sim \alpha(\cdot)$ rassal değişkenimiz olsun.

$Z = z$ verildiğinde $X|Z = z \sim p_z(\cdot)$ olsun.

Bu durumda, X'in marginal (tekil) dağılımı bir *karışım dağılımıdır*.

$$\pi(x) = \begin{cases} \int p_z(x)\alpha(z)dz, & \alpha(z) \text{ olasılık yoğunluk fonksiyonu ise} \\ \sum_z p_z(x)\alpha(z), & \alpha(z) \text{ olasılık kütle fonksiyonu ise} \end{cases}$$

$X \sim \pi$ nasıl üretilebilir?

Birleştirme yöntemi

$X \sim \pi$ nasıl üretilebilir?

$$\pi(x) = \begin{cases} \int p_z(x)\alpha(z)dz, & \alpha(z) \text{ olasılık yoğunluk fonksiyonu ise} \\ \sum_z p_z(x)\alpha(z), & \alpha(z) \text{ olasılık kütle fonksiyonu ise} \end{cases}$$

Doğrudan π'den örnekleme çok zor olabilir, ancak α ve p_z'den örnekleme yapmak kolaysa, birleştirme yöntemi kullanılabilir:

1. $Z \sim \alpha(\cdot)$ üretilir,
2. $X \sim p_Z(\cdot)$, üretilir
3. Z atılır ve X tutulur.

Bu şekilde üretilen X kesin olarak π'den gelir.

Örnek: Karışım Gauss dağılımı

K bileşeni olan, bileşenlerinin

- ortalama değerleri ve varyansları: $(\mu_1, \sigma_1^2), \ldots, (\mu_K, \sigma_K^2)$
- karışımdaki olasılık ağırlıkları $w_1, \ldots, w_K$ $(w_1 + \cdots + w_K = 1)$

olan karışım Gauss dağılımının yoğunluk fonksiyonu

$$\pi(x) = \sum_{k=1}^{K} w_k \phi(x; \mu_k, \sigma_k^2).$$

Bu dağılımdan örnekleme yapmak için

1. w_k olasılıkla k üretilir,
2. $X \sim \mathcal{N}(\mu_k, \sigma_k^2)$ üretilir,
3. k atılır ve X tutulur.

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Reddetme örneklemesi

Reddetme örneklemesi

Sık kullanılan bir başka yöntem.

Şu şartları sağlayan bir $q(x)$ dağılımı gerekir.

- $\pi(x) > 0$ ise $q(x) > 0$ olmalı
- Her $x \in \mathcal{X}$ için $\pi(x) \leq Mq(x)$'yi sağlayacak bir $M > 0$ olması.

Reddetme örneklemesi:

1. $X' \sim q(\cdot)$ ve $U \sim \text{Unif}(0, 1)$ üretilir.
2. $U \leq \frac{\pi(X')}{Mq(X')}$, ise $X = X'$ alınır; değilse 1.'e geri dönülür.

Reddetme örneklemesi: Kabul olasılığı

1. $X' \sim q(\cdot)$ ve $U \sim \text{Unif}(0,1)$ üretilir.
2. $U \leq \frac{\pi(X')}{Mq(X')}$, ise $X = X'$ alınır; değilse 1.'e geri dönülür.

Bir döngüde kabul etme olasılığı

$$
\begin{aligned}
\mathbb{P}(\text{Kabul}) &= \int \mathbb{P}(\text{Kabul}|X' = x) p_{X'}(x) dx \\
&= \int \frac{\pi(x)}{Mq(x)} q(x) dx \\
&= \frac{1}{M} \int \pi(x) dx \\
&= \frac{1}{M}.
\end{aligned}
$$

Dolayısıyla $q(x)$'i $\pi(x)$'e olabildiğince yakın seçmek ve $M = \sup_x \pi(x)/q(x)$ almak makuldür.

$\pi(x)$ tam olarak bilinmediğinde

Diyelim ki $\pi(x)$'in sadece bilinmeyen bir sabit çarpanla çarpılmış haldeki değerini biliyoruz:

$$\pi(x) = \frac{\widehat{\pi}(x)}{Z_\pi}, \quad Z_\pi = \int \widehat{\pi}(x) dx$$

Reddetme örneklemesi, bütün $x \in \mathcal{X}$ için $\widehat{\pi}(x) \leq Mq(x)$'i sağlayan bir M ile hala uygulanabilir.

1. $X' \sim q(\cdot)$ ve $U \sim \text{Unif}(0,1)$ üretilir.
2. $U \leq \frac{\widehat{\pi}(X')}{Mq(X')}$ ise, $X = X'$ alınır; değilse 1.'e gidilir.

Kabul olasılığı: $\frac{1}{M} Z_\pi$.

$\pi(x)$ tam olarak bilinmediğinde: Bayesci çıkarım örneği

Bilinmeyen sabit sorunu Bayesci çıkarımda sıklıkla karşımıza çıkar.

Bayesci çıkarımda amaç sonsal dağılımı bulmaktır.

Hesaplanamayan sonsal dağılımlardan örnekleme yapılabilir.

X'in $Y = y$ verildiğindeki sonsal dağılımı

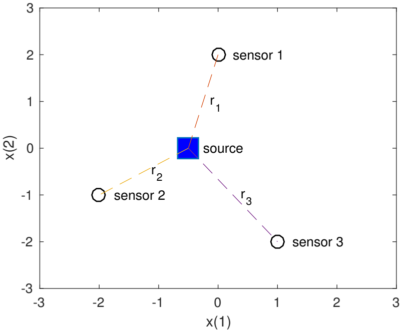
$$\pi(x) := p_{X|Y}(x|y) \propto p_X(x)p_{Y|X}(y|x) = \widehat{\pi}(x)$$

Çarpımsal (ve çoğunlukla hesaplanamayan) sabit:

$$p_Y(y) = \int p_X(x)p_{Y|X}(y|x)dx$$

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatlarını saptamak istediğimiz bir hedef (source) $X = (X(1), X(2))$:



s_1, s_2, s_3 noktalarındaki üç sensör hedefe olan uzaklıklarını ölçüyor:

$$r_i = [(X(1) - s_i(1))^2 + (X(2) - s_i(2))^2]^{1/2}, \quad i = 1, 2, 3$$

Ölçümler $Y = (Y_1, Y_2, Y_3)$ gürültülü:

$$Y_i = r_i + V_i, \quad V_i \sim \mathcal{N}(0, \sigma_y^2), \quad i = 1, 2, 3.$$

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatla ilgili önsel kanı:

$$X \sim \mathcal{N}(0_2, \sigma_x^2 I_2), \quad \sigma_x^2 \gg 1$$

Amaç: $Y = y = (y_1, y_2, y_3)$ verildiğinde $\mathbb{E}(X|Y = y)$'yi hesaplamak.

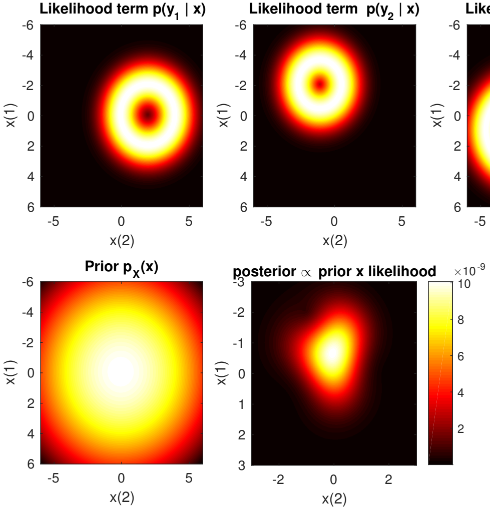
$$\pi(x) := p_{X|Y}(x|y) \propto \underbrace{p_X(x) p_{Y|X}(y|x)}_{\widehat{\pi}(x)} = \underbrace{\phi(x; 0_2, \sigma_x^2 I_2)}_{p_X(x)} \underbrace{\prod_{i=1}^{3} \phi(y_i; r_i, \sigma_y^2)}_{p_{Y|X}(y|x)}$$

Reddetme örneklemesi $q(x) = p_X(x)$ alınarak yapılabilir:

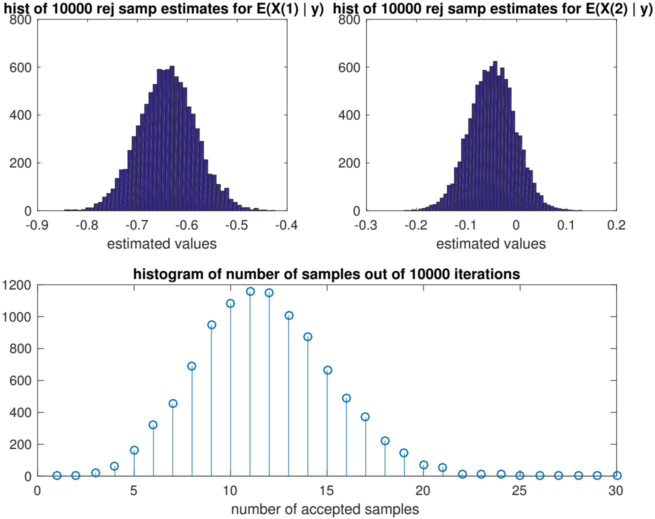
$$\frac{\widehat{\pi}(x)}{q(x)} = p_{Y|X}(y|x) = \prod_{i=1}^{3} \phi(y_i; r_i, \sigma_y^2) = \frac{1}{(2\pi\sigma_y^2)^{3/2}} e^{-\frac{1}{2}\sum_{i=1}^{3}(y_i - r_i)^2} \leq \frac{1}{(2\pi\sigma_y^2)^{3/2}}$$

O halde, $M = \frac{1}{(2\pi\sigma_y^2)^{3/2}}$ seçilmelidir.

Hedef yer saptaması - dağılımlar, $\sigma_x^2 = 100$, $\sigma_y^2 = 1$



Hedef yer saptaması - reddetme örn. $\sigma_x^2 = 100$, $\sigma_y^2 = 1$



Markov zinciri Monte Carlo

Ayrık zamanlı Markov zinciri

Başlangıç yoğunluk/kütle fonksiyonu ve geçiş olasılık çekirdeği yoğunluk/kütle fonksiyonu sırasıyla $\eta(x)$ ve $M(x'|x)$ olan bir Markov zinciri $\{X_t\}_{t\geq}1$:

$$\begin{aligned} p(x_{1:n}) &= \eta(x_1)M(x_2|x_1)\ldots M(x_n|x_{n-1}) \\ &= \eta(x_1)\prod_{t=2}^{n} M(x_t|x_{t-1}) \end{aligned}$$

Geçmiş değerler verildiğinde, Markov zincirinin n zamanındaki değeri sadece $n-1$ zamandaki değerine bağlıdır.

$$\begin{aligned} p(x_n|x_{1:n-1}) &= p(x_n|x_{n-1}) \\ &= M(x_n|x_{n-1}). \end{aligned}$$

Değişimsiz dağılım ve durağan dağılım

X_n'in marjinal dağılımını özyinelemeli olarak yazabiliriz

$$\pi_1(x) := \eta(x)$$
$$\pi_n(x) := \int M(x|x')\pi_{n-1}(x')dx'$$

Eğer verilen bir $\pi(x)$ dağılımı

$$\pi(x) = \int M(x|x')\pi(x')dx'$$

şartını sağlıyorsa "$\pi(x)$, M'ye göre değişimsizdir" denir ve M'nin belli şartları sağlaması durumunda

- $\pi(x)$, M'nin tek değişimsiz dağılımıdır,
- $\pi(x)$, M'nin durağan dağılımıdır, yani

$$\lim_{n\to\infty} \pi_n \to \pi$$

Markov zinciri Monte Carlo

Örnekleme problemi: $\pi(x)$ dağılımından örnekleme yapmak.

Markov zinciri Monte Carlo (MZMC) yöntemleri, durağan dağılımı π olan bir Markov zincirinin tasarımına dayanır.

Bu zincir yeterince uzun zaman çalıştırıldığında (mesela bir t_b zamanından sonra) zincirin üretilen değerlerinin $X_{t_b+1}, X_{t_b+2}, \ldots, X_T$ yaklaşık olarak π'den geldiği kabul edilir.

Bu değerler, π dağılıma göre olan beklenti değerlerini hesaplamaya yarar.

$$\pi(\varphi) \approx \frac{1}{T - t_b} \sum_{t=t_b+1}^{T} \varphi(X_t)$$

Gibbs örneklemesi

Gibbs örneklemesi

En çok kullanılan MZMC yöntemlerinden biri *Gibbs örneklemesidir*.

Uygulanması için

- $X = (X(1), \ldots, X(d))$ değişkeni çok boyutlu olmalı,
- tam koşullu $\pi_k\left(\cdot | X(1), \ldots X(k-1), X(k+1), \ldots, X(d)\right)$ dağılımlarından örnekleme yapılabilmeli.

Gibbs örneklemesi:

$X_1 = (X_1(1), \ldots, X_1(d))$ ile başla.

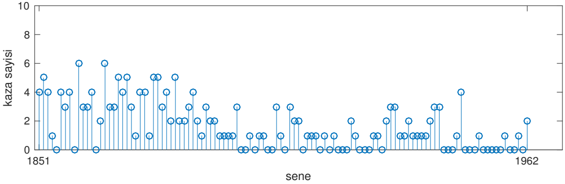
$n = 2, 3, \ldots$ için,

$k = 1, \ldots, d$ için

$$X_n(k) \sim \pi_k(\cdot | X_n(1), \ldots, X_n(k-1), X_{n-1}(k+1), \ldots, X_{n-1}(n)).$$

Değişim noktası modelleri

İngiltere'deki kömür madenlerinde 1851-1962 arasında meydana gelen kaza sayıları:



Kaza sürecini bir heterojen Poisson süreci olarak düşünebiliriz.

Soru: Bu yıllar boyunca kaza sıklığında bir 'değişim' olmuş ise bu ne zaman olmuş?

$$Y_t \sim \begin{cases} \mathcal{PO}(\lambda_1), & 1 \leq t \leq \tau \\ \mathcal{PO}(\lambda_2), & \tau < t \leq n. \end{cases}$$

Bilinmeyen parametreler: $x = (\tau, \lambda_1, \lambda_2)$. Öncül dağılımlar:

$$\lambda_i \sim \text{Gamma}(\alpha, \beta), \quad i = 1, 2, \quad \tau \sim \text{Unif}\{1, \dots, n\}.$$

Değişim noktası modelleri - Gibbs örnekleyicisi

Bileşik dağılım fonksiyonu:

$$\begin{aligned} p(\tau, \lambda_1, \lambda_2 | y_{1:n}) &\propto p(\tau, \lambda_1, \lambda_2, y_{1:n}) \\ &= p(\tau) p(\lambda_1) p(\lambda_2) p(y_{1:n} | \tau, \lambda_1, \lambda_2) \\ &= \frac{1}{n} \frac{\beta^\alpha \lambda_1^{\alpha-1} e^{-\beta\lambda_1}}{\Gamma(\alpha)} \frac{\beta^\alpha \lambda_2^{\alpha-1} e^{-\beta\lambda_2}}{\Gamma(\alpha)} \prod_{t=1}^{\tau} \frac{e^{-\lambda_1} \lambda_1^{y_t}}{y_t!} \prod_{t=\tau+1}^{n} \frac{e^{-\lambda_2} \lambda_2^{y_t}}{y_t!} \end{aligned}$$

Koşullu dağılımlar:

$$\lambda_1 | \tau, \lambda_2, y_{1:n} \sim \text{Gamma}\left(\alpha + \sum_{t=1}^{\tau} y_t, \beta + \tau \right)$$

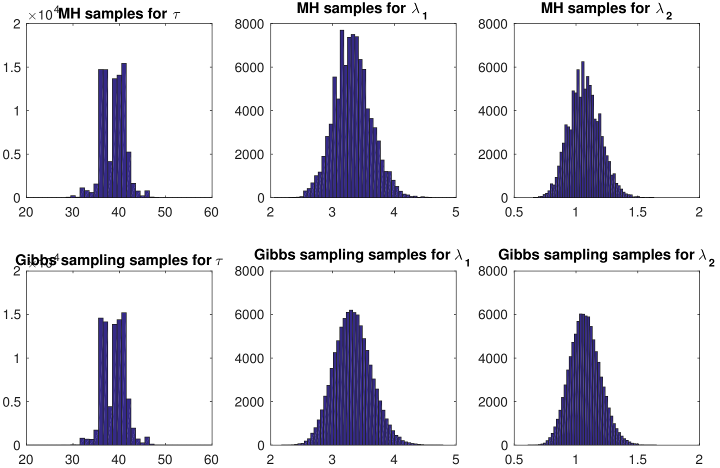
$$\lambda_2 | \tau, \lambda_1, y_{1:n} \sim \text{Gamma}\left(\alpha + \sum_{t=\tau+1}^{n} y_t, \beta + n - \tau \right)$$

$$\tau | \lambda_1, \lambda_2, y_{1:n} \sim \text{Categorical}(a_1, \ldots, a_n)$$

$$a_i = \frac{e^{-i\lambda_1} \lambda_1^{\sum_{t=1}^{i} y_t} e^{-(n-i)\lambda_2} \lambda_2^{\sum_{t=i+1}^{n} y_t}}{\sum_{j=1}^{n} \left[e^{-j\lambda_1} \lambda_1^{\sum_{t=1}^{j} y_t} e^{-(n-j)\lambda_2} \lambda_2^{\sum_{t=j+1}^{n} y_t} \right]}$$

Değişim noktası modelleri - sonuçlar

MH ve Gibbs örneklemesi 100000 döngü boyunca çalıştırıldı.



Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Metropolis-Hastings

Metropolis-Hastings

Bir diğer sık kullanılan MZMC yöntemi de *Metropolis-Hastings* yöntemidir.

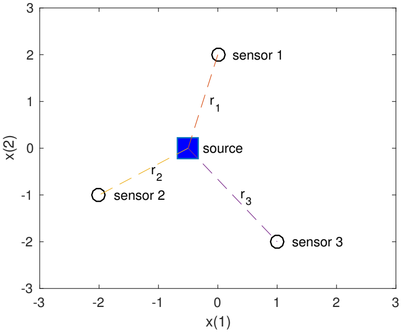
$X_{n-1} = x$ verildiğinde yeni değer için $q(\cdot|x)$ koşullu dağılımından çekilen bir örnek yeni değer olarak önerilir, bu değer belli bir olasılığa göre kabul edilir, edilmezse eski değerde kalınır.

$X_{n-1} = x$ verildiğinde,

- Yeni değer için $x' \sim q(\cdot|x)$ önerilir.
- X_n'in değeri
$$\alpha(x, x') = \min\left\{1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)}\right\}$$
olasılıkla $X_n = x'$ alınır, yoksa önerilen değer reddedilir ve $X_n = x$ alınır.

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatlarını saptamak istediğimiz bir hedef (source) $X = (X(1), X(2))$:



s_1, s_2, s_3 noktalarındaki üç sensör hedefe olan uzaklıklarını ölçüyor:

$$r_i = [(X(1) - s_i(1))^2 + (X(2) - s_i(2))^2]^{1/2}, \quad i = 1, 2, 3$$

Ölçümler $Y = (Y_1, Y_2, Y_3)$ gürültülü:

$$Y_i = r_i + V_i, \quad V_i \sim \mathcal{N}(0, \sigma_y^2), \quad i = 1, 2, 3.$$

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatla ilgili önsel kanı:

$$X \sim \mathcal{N}(0_2, \sigma_x^2 I_2), \quad \sigma_x^2 \gg 1$$

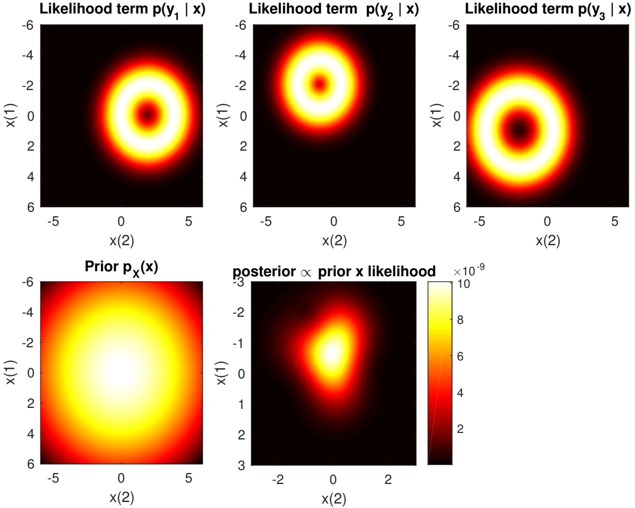
Amaç: $Y = y = (y_1, y_2, y_3)$ verildiğinde $p_{X|Y}(x|y)$'i bulmak ve $\mathbb{E}(X|Y = y)$'yi hesaplamak.

$$\pi(x) = p_{X|Y}(x|y) \propto \underbrace{p_X(x) p_{Y|X}(y|x)}_{\widehat{\pi}(x)} = \underbrace{\phi(x; 0_2, \sigma_x^2 I_2)}_{p_X(x)} \underbrace{\prod_{i=1}^{3} \phi(y_i; r_i, \sigma_y^2)}_{p_{Y|X}(y|x)}$$

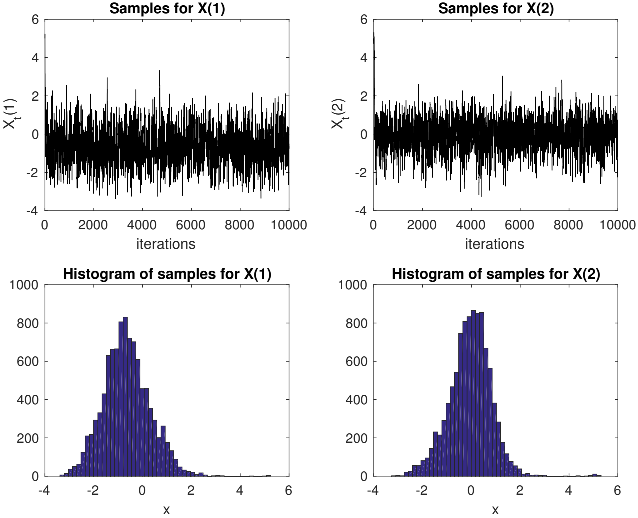
Metropolis-Hastings $q(x'|x) = \phi(x'; x, \sigma_q^2 I_2)$ alınarak çalıştırılabilir:

$$\alpha(x, x') = \min\left\{1, \frac{p_X(x') p_{Y|X}(y|x') q(x|x')}{p_X(x) p_{Y|X}(y|x) q(x'|x)}\right\} = \min\left\{1, \frac{p_X(x') p_{Y|X}(y|x')}{p_X(x) p_{Y|X}(y|x)}\right\}$$

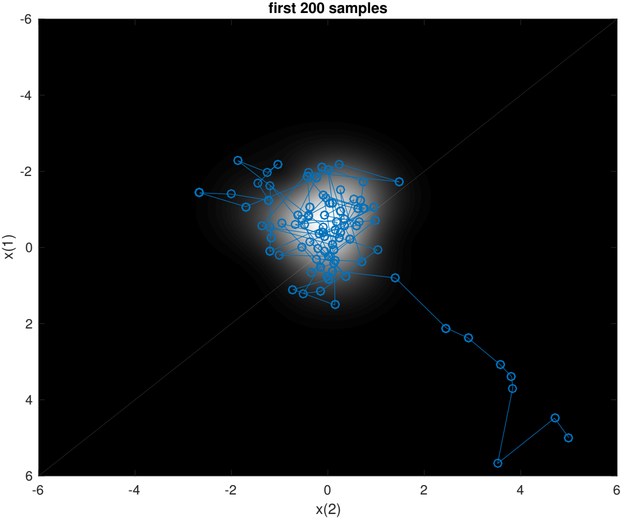
Hedef yer saptaması - dağılımlar, $\sigma_x^2 = 100$, $\sigma_y^2 = 1$



Hedef yer saptaması için Metropolis-Hastings: örnekler



Hedef yer saptaması için Metropolis-Hastings: ilk 200 örnek



Lojistik regresyon

```
25, Private, 226802, 11th, 7, Never-married, Machine-op-inspct, Own-child, Black, Male, 0, 0, 40, United-States, <=50K.
38, Private, 89814, HS-grad, 9, Married-civ-spouse, Farming-fishing, Husband, White, Male, 0, 0, 50, United-States, <=50K.
28, Local-gov, 336951, Assoc-acdm, 12, Married-civ-spouse, Protective-serv, Husband, White, Male, 0, 0, 40, United-States, >50K.
44, Private, 160323, Some-college, 10, Married-civ-spouse, Machine-op-inspct, Husband, Black, Male, 7688, 0, 40, United-States, >50K.
18, ?, 103497, Some-college, 10, Never-married, ?, Own-child, White, Female, 0, 0, 30, United-States, <=50K.
34, Private, 198693, 10th, 6, Never-married, Other-service, Not-in-family, White, Male, 0, 0, 30, United-States, <=50K.
29, ?, 227026, HS-grad, 9, Never-married, ?, Unmarried, Black, Male, 0, 0, 40, United-States, <=50K.
63, Self-emp-not-inc, 104626, Prof-school, 15, Married-civ-spouse, Prof-specialty, Husband, White, Male, 3103, 0, 32, United-States,
24, Private, 369667, Some-college, 10, Never-married, Other-service, Unmarried, White, Female, 0, 0, 40, United-States, <=50K.
55, Private, 104996, 7th-8th, 4, Married-civ-spouse, Craft-repair, Husband, White, Male, 0, 0, 10, United-States, <=50K.
65, Private, 184454, HS-grad, 9, Married-civ-spouse, Machine-op-inspct, Husband, White, Male, 6418, 0, 40, United-States, >50K.
36, Federal-gov, 212465, Bachelors, 13, Married-civ-spouse, Adm-clerical, Husband, White, Male, 0, 0, 40, United-States, <=50K.
26, Private, 82091, HS-grad, 9, Never-married, Adm-clerical, Not-in-family, White, Female, 0, 0, 39, United-States, <=50K.
58, ?, 299831, HS-grad, 9, Married-civ-spouse, ?, Husband, White, Male, 0, 0, 35, United-States, <=50K.
48, Private, 279724, HS-grad, 9, Married-civ-spouse, Machine-op-inspct, Husband, White, Male, 3103, 0, 48, United-States, >50K.
43, Private, 346189, Masters, 14, Married-civ-spouse, Exec-managerial, Husband, White, Male, 0, 0, 50, United-States, >50K.
20, State-gov, 444554, Some-college, 10, Never-married, Other-service, Own-child, White, Male, 0, 0, 25, United-States, <=50K.
43, Private, 128354, HS-grad, 9, Married-civ-spouse, Adm-clerical, Wife, White, Female, 0, 0, 30, United-States, <=50K.
37, Private, 60548, HS-grad, 9, Widowed, Machine-op-inspct, Unmarried, White, Female, 0, 0, 20, United-States, <=50K.
40, Private, 85019, Doctorate, 16, Married-civ-spouse, Prof-specialty, Husband, Asian-Pac-Islander, Male, 0, 0, 45, ?, >50K.
34, Private, 107914, Bachelors, 13, Married-civ-spouse, Tech-support, Husband, White, Male, 0, 0, 47, United-States, >50K.
34, Private, 238588, Some-college, 10, Never-married, Other-service, Own-child, Black, Female, 0, 0, 35, United-States, <=50K.
72, ?, 132015, 7th-8th, 4, Divorced, ?, Not-in-family, White, Female, 0, 0, 6, United-States, <=50K.
25, Private, 220931, Bachelors, 13, Never-married, Prof-specialty, Not-in-family, White, Male, 0, 0, 43, Peru, <=50K.
25, Private, 205947, Bachelors, 13, Married-civ-spouse, Prof-specialty, Husband, White, Male, 0, 0, 40, United-States, <=50K.
45, Self-emp-not-inc, 432824, HS-grad, 9, Married-civ-spouse, Craft-repair, Husband, White, Male, 7298, 0, 90, United-States, >50K.
22, Private, 236427, HS-grad, 9, Never-married, Adm-clerical, Own-child, White, Male, 0, 0, 20, United-States, <=50K.
23, Private, 134446, HS-grad, 9, Separated, Machine-op-inspct, Unmarried, Black, Male, 0, 0, 54, United-States, <=50K.
54, Private, 99516, HS-grad, 9, Married-civ-spouse, Craft-repair, Husband, White, Male, 0, 0, 35, United-States, <=50K.
32, Self-emp-not-inc, 109282, Some-college, 10, Never-married, Prof-specialty, Not-in-family, White, Male, 0, 0, 60, United-States, <
46, State-gov, 106444, Some-college, 10, Married-civ-spouse, Exec-managerial, Husband, Black, Male, 7688, 0, 38, United-States, >50K.
56, Self-emp-not-inc, 186651, 11th, 7, Widowed, Other-service, Unmarried, White, Female, 0, 0, 50, United-States, <=50K.
24, Self-emp-not-inc, 188274, Bachelors, 13, Never-married, Sales, Not-in-family, White, Male, 0, 0, 50, United-States, <=50K.
23, Local-gov, 258120, Some-college, 10, Married-civ-spouse, Protective-serv, Husband, White, Male, 0, 0, 40, United-States, <=50K.
26, Private, 43311, HS-grad, 9, Divorced, Exec-managerial, Unmarried, White, Female, 0, 0, 40, United-States, <=50K.
65, ?, 191846, HS-grad, 9, Married-civ-spouse, ?, Husband, White, Male, 0, 0, 40, United-States, <=50K.
36, Local-gov, 403681, Bachelors, 13, Married-civ-spouse, Prof-specialty, Husband, White, Male, 0, 0, 40, United-States, >50K.
22, Private, 248446, 5th-6th, 3, Never-married, Priv-house-serv, Not-in-family, White, Male, 0, 0, 50, Guatemala, <=50K.
17, Private, 269430, 10th, 6, Never-married, Machine-op-inspct, Not-in-family, White, Male, 0, 0, 40, United-States, <=50K.
20, Private, 257509, HS-grad, 9, Never-married, Craft-repair, Own-child, White, Male, 0, 0, 40, United-States, <=50K.
65, Private, 136384, Masters, 14, Married-civ-spouse, Prof-specialty, Husband, White, Male, 0, 0, 50, United-States, >50K.
```

Lojistik Regresyon

- Gözlenenler:
  - Öznitelikler:

$$z_t \in \mathbb{R}^d, \quad t = 1, \ldots, n$$

  - İkili cevap:

$$Y_t \in \{-1, +1\}, \quad t = 1, \ldots, n$$

- Bilinmeyen: Y_t ile z_t arasındaki ilişki.

Lojistik regresyon: İlişki bir x parametresiyle modellenir.

$$\mathbb{P}(Y_t = y_t | z_t, x) = \frac{1}{1 + e^{-y_t z_t^T x}}, \quad y_t = -1, 1,$$

Problem: $\{(y_t, z_t), t = 1, \ldots, n\}$ verildiğinde x''in kestirimi.

Bayesci yaklaşım: $p(x|y_{1:n}, z_{1:n})$'nin hesaplanması.

Önem örneklemesi

Önem örneklemesi: Motivasyon

Yola çıkarkenki problemimiz: yaklaşık hesaplamak istediğimiz beklenti değeri

$$\pi(\varphi) = \mathbb{E}_\pi(\varphi(X)) = \int_{\mathcal{X}} \varphi(x)\pi(x)dx.$$

$\pi(\varphi)$'nin Monte Carlo kestirimi

$$\pi^N_{\mathsf{MC}}(\varphi) = \frac{1}{N}\sum_{i=1}^{N}\varphi(X^{(i)}), \quad X^{(i)} \sim \pi, \quad i = 1, \dots, N,$$

için π'den örnekleme yapmamız gerekiyor.

Bir çok durumda $X \sim \pi$ örneklemesi çok zor, çok pahalı veya imkansız olabilir.

Önem örneklemesi

Yine, $\pi(x) > 0$ ise $q(x) > 0$ şartını sağlayan bir yardımcı dağılımımız olsun. $\pi(x)$ ve $q(x)$ verildiğinde, önem fonksiyonunu tanımlayalım $w : \mathcal{X} \to \mathbb{R}$

$$w(x) := \begin{cases} \pi(x)/q(x), & q(x) > 0, \\ 0 & q(x) = 0. \end{cases}$$

Önem örneklemesine temel oluşturan bağlantı:

$$\begin{aligned} \pi(\varphi) = \mathbb{E}_\pi(\varphi(X)) &= \int_\mathcal{X} \varphi(x)\pi(x)dx \\ &= \int_\mathcal{X} \varphi(x)\frac{\pi(x)}{q(x)}q(x)dx \\ &= \int_\mathcal{X} \varphi(x)w(x)q(x)dx \\ &= \mathbb{E}_q(\varphi(X)w(X)) \end{aligned}$$

Önem örneklemesi

Eğer $q(x)$'ten örnekleme yapmak kolay ise, $\pi(\varphi)$'ye yaklaşmak için önem örneklemesi yapılabilir.

1. $X^{(1)}, \ldots, X^{(N)} \overset{\text{i.i.d.}}{\sim} q(\cdot)$ örneklenir.
2. $\pi(\varphi)$'ye şu şekilde yaklaşılır:

$$\pi_{\text{IS}}^{N}(\varphi) := \frac{1}{N} \sum_{i=1}^{N} \varphi(X^{(i)}) w(X^{(i)}).$$

Öz-düzgeleyici önem örneklemesi

Önem örneklemesi $\pi(x) = \frac{\widehat{\pi}(x)}{Z_\pi}$ olduğunda ve sadece $\widehat{\pi}(x)$ bilindiğinde yine uygulanabilir.

Önem fonksiyonu

$$w(x) := \begin{cases} \frac{\widehat{\pi}(x)}{q(x)}, & q(x) > 0 \\ 0, & \widehat{q}(x) = 0, \end{cases}$$

$$\mathbb{E}(w(X)) = \int \frac{\widehat{\pi}(x)}{q(x)} q(x) dx = \int \frac{\pi(x) Z_\pi}{q(x)} q(x) dx = Z_\pi.$$

$$\mathbb{E}(w(X)\varphi(X)) = \int \frac{\widehat{\pi}(x)}{q(x)} \varphi(x) q(x) dx = \int \frac{\pi(x) Z_\pi}{q(x)} \varphi(x) q(x) dx = \pi(\varphi) Z_\pi.$$

İki ifadeyi birbirine bölersek

$$\pi(\varphi) = \frac{\mathbb{E}(w(X)\varphi(X))}{Z_\pi} = \frac{\mathbb{E}(w(X)\varphi(X))}{\mathbb{E}(w(X))}.$$

Öz-düzgeleyici önem örneklemesi

$$\pi(\varphi) = \frac{\mathbb{E}(w(X)\varphi(X))}{Z_\pi} = \frac{\mathbb{E}(w(X)\varphi(X))}{\mathbb{E}(w(X))}.$$

Hem pay hem payda için aynı örnekler kullanarak önem örneklemesi yapılabilir.

$$\pi_{\mathsf{IS}}^N(\varphi) = \frac{\frac{1}{N}\sum_{i=1}^N \varphi(X^{(i)})w(X^{(i)})}{\frac{1}{N}\sum_{i=1}^N w(X^{(i)})}, \quad X^{(1)}, \ldots, X^{(N)} \sim q(\cdot).$$

Öz-düzgeleyici önem ağırlıkları:

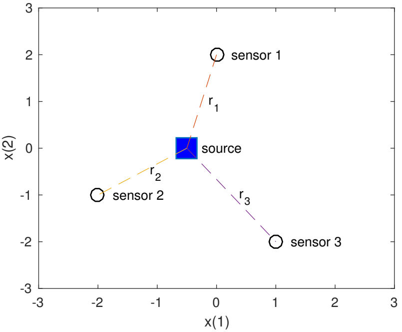
$$W^{(i)} = \frac{w(X^{(i)})}{\sum_{j=1}^N w(X^{(j)})}$$

Öz-düzgeleyici önem örneklemesi

1. $i = 1, \ldots, N$ için; $X^{(i)} \sim q(\cdot)$ üretilir ve $w(X^{(i)}) = \frac{\widehat{\pi}(X^{(i)})}{q(X^{(i)})}$ hesaplanır.
2. $i = 1, \ldots, N$ için $W^{(i)} = \frac{w(X^{(i)})}{\sum_{j=1}^N w(X^{(j)})}$ hesaplanır.
3. $\pi_{\mathsf{IS}}^N(\varphi) = \sum_{i=1}^N W^{(i)}\varphi(X^{(i)})$ hesaplanır.

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatlarını saptamak istediğimiz bir hedef (source) $X = (X(1), X(2))$:



s_1, s_2, s_3 noktalarındaki üç sensör hedefe olan uzaklıklarını ölçüyor:

$$r_i = [(X(1) - s_i(1))^2 + (X(2) - s_i(2))^2]^{1/2}, \quad i = 1, 2, 3$$

Ölçümler $Y = (Y_1, Y_2, Y_3)$ gürültülü:

$$Y_i = r_i + V_i, \quad V_i \sim \mathcal{N}(0, \sigma_y^2), \quad i = 1, 2, 3.$$

Bayesci çıkarım örneği: Hedef yer saptaması

Koordinatla ilgili önsel kanı:

$$X \sim \mathcal{N}(0_2, \sigma_x^2 I_2), \quad \sigma_x^2 \gg 1$$

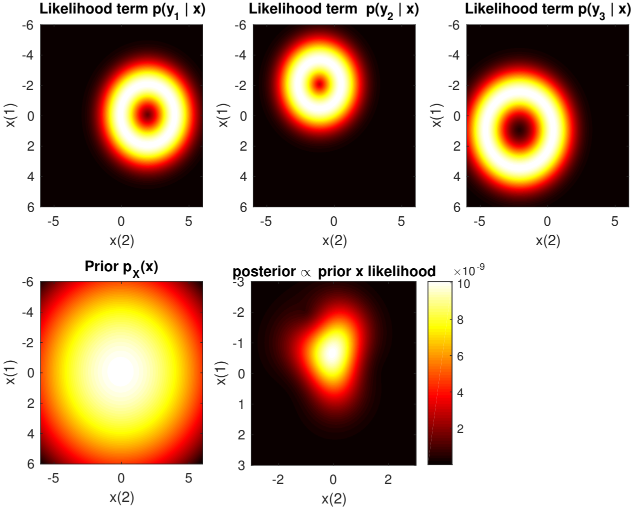
Amaç: $Y = y = (y_1, y_2, y_3)$ verildiğinde $p_{X|Y}(x|y)$'i bulmak ve $\mathbb{E}(X|Y = y)$'yi hesaplamak.

$$\pi(x) = p_{X|Y}(x|y) \propto \underbrace{p_X(x) p_{Y|X}(y|x)}_{\widehat{\pi}(x)} = \underbrace{\phi(x; 0_2, \sigma_x^2 I_2)}_{p_X(x)} \underbrace{\prod_{i=1}^{3} \phi(y_i; r_i, \sigma_y^2)}_{p_{Y|X}(y|x)}$$

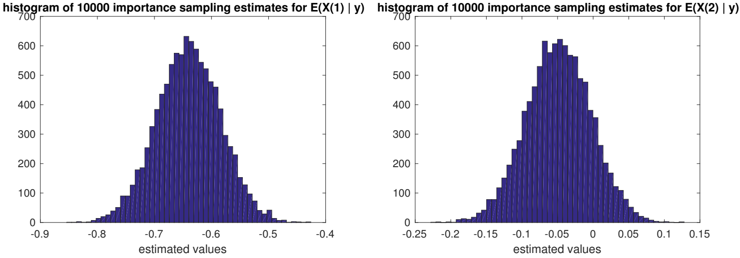
Öz-düzgeleyici önem örneklemesi $q(x) = p_X(x)$ alınarak yapılabilir:

$$w(x) = \frac{\widehat{\pi}(x)}{q(x)} = \frac{p_X(x) p_{Y|X}(y|x)}{p_X(x)} = p_{Y|X}(y|x) = \prod_{i=1}^{3} \phi(y_i; r_i, \sigma_y^2)$$

Hedef yer saptaması - dağılımlar, $\sigma_x^2 = 100$, $\sigma_y^2 = 1$



Hedef yer saptaması



Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Sıralı Monte Carlo

Büyüyen boyutlarda sıralı çıkarım

$\{X_n\}_{n \geq 1}$, her biri $\mathcal{X}$'ten değer alan rassal değişkenler dizisi olsun.

$X_{1:n}$ için $\{\pi_n(x_{1:n})\}_{n \geq 1}$ dağılım dizisi verilmiş olsun.

Her biri $\varphi_n : X^n \to \mathbb{R}$ olan bir $\{\varphi_n\}_{n \geq 1}$ fonksiyon dizisi verilsin.

Amaç: Sıralı çıkarım

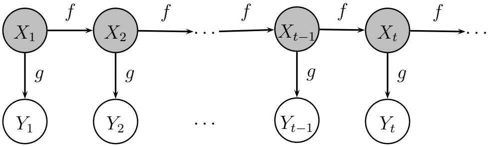
$$\pi_n(\varphi_n) = \mathbb{E}_{\pi_n}\left[\varphi_n(X_{1:n})\right] = \int \pi_n(x_{1:n})\varphi_n(x_{1:n})dx_{1:n}, \quad n = 1, 2, \ldots$$

integrallerine sıralı bir şekilde nasıl yaklaşabiliriz?

Örnek: Saklı Markov modelleri (SMM)

SMM, biri gizli ve Markov zinciri olan, diğeri gözlenen iki süreçten oluşur.

$$\{X_t \in \mathcal{X}, Y_t \in \mathcal{Y}\}_{t \geq 1}$$



1. $\{X_t\}_{t\geq 1}$ başlangıç ve geçiş yoğunlukları $\eta(x)$ ve $f(x'|x)$ olan saklı Markov zinciri

$$X_1 \sim \eta(x), \quad X_t|(X_{1:t-1} = x_{1:t-1}) \sim f(\cdot|x_{t-1}),$$

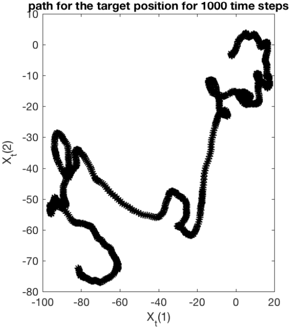
2. $\{Y_t\}_{t\geq 1}$, $\{X_t\}_{t\geq 1}$'ye koşullu bağımsız süreç:

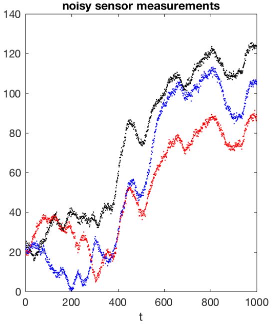
$$Y_t|(\{X_i\}_{i\geq 1} = \{x_i\}_{i\geq 1}, \{Y_i\}_{i\neq t} = \{y_i\}_{i\neq t}) \sim g(\cdot|x_t).$$

Örneğe örnek: Hedef takip

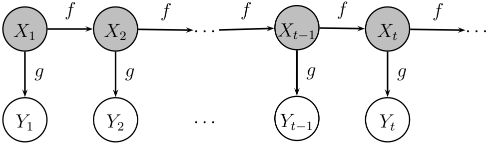
1. $X_t = (V_t, P_t)$: t anındaki hız ve pozisyon
   - $V_t = (V_t(1), V_t(2))$: t anındaki hız vektörü
   - $P_t = (P_t(1), P_t(2))$: t anındaki pozisyon vektörü
2. $Y_t \sim \mathcal{N}((||S_1 - P_t||, ||S_2 - P_t||, ||S_3 - P_t||), \sigma_y^2 I_3)$: 3 sensörlerden alınan gürültülü uzaklık ölçümleri.

X_t bir Markov zinciri olarak modellenebilir. Bu durumda $\{X_t, Y_t\}$ bir SMM oluşturur.





SMM: hedef sonsal dağılımlar



Ortak dağlım:

$$p(x_{1:n}, y_{1:n}) = \eta(x_1) \prod_{t=2}^{n} f(x_t|x_{t-1}) \prod_{t=1}^{n} g(y_t|x_t)$$

Gözlemlerin marjinal (tekil) dağılımı

$$p(y_{1:n}) = \int_{\mathcal{X}^n} p(x_{1:n}, y_{1:n}) dx_{1:n}.$$

$x_{1:n}$'nin $y_{1:n}$'e olan sonsal dağılımı:

$$p(x_{1:n}|y_{1:n}) = \frac{p(x_{1:n}, y_{1:n})}{p(y_{1:n})} \propto p(x_{1:n}, y_{1:n})$$

Amaç: $\pi_n(x_{1:n}) = p(x_{1:n}|y_{1:n})$ ve $\mathbb{E}_{\pi_n}[\varphi_n(X_{1:n})]$'e yaklaşmak.

Sıralı önem örneklemesi

Sıralı önem örneklemesi

$\mathbb{E}_{\pi_n}[\varphi_n(X_{1:n})]$ için önem örneklemesi yapmak istiyoruz.

Bunun için $q_n(x_{1:n})$'ye ihtiyacımız var, bu durumda ağırlık fonksiyonları

$$w_n(x_{1:n}) = \frac{\pi_n(x_{1:n})}{q_n(x_{1:n})}.$$

q_n'yi sıralı olarak oluşturabiliriz:

$$q_n(x_{1:n}) = q_1(x_1) \prod_{i=1}^{n} q_i(x_i | x_{1:i-1})$$

Bu durumda ağırlık fonksiyonları özyinelemeli olarak yazılabilir:

$$w_n(x_{1:n}) = w_{n-1}(x_{1:n-1}) \underbrace{\frac{\pi_n(x_{1:n})}{\pi_{n-1}(x_{1:n-1}) q_n(x_n | x_{1:n-1})}}_{w_{n|n-1}(x_{1:n})}.$$

$\pi_n(x_{1:n}) = \widehat{\pi}_n(x_{1:n})/Z_n$ ve $\widehat{\pi}_n(x_{1:n})$ biliniyorsa,

$$W_n^{(i)} = \frac{w_n(X_{1:n}^{(i)})}{\sum_{i=1}^{N} w_n(X_{1:n}^{(i)})}.$$

Sıralı (öz-düzgeleyici) önem örneklemesi

Diyelim ki $\pi_n(x_{1:n}) = \widehat{\pi}_n(x_{1:n})/Z_n$ ve $\widehat{\pi}_n(x_{1:n})$ biliniyor.

Öz-düzgeleyici önem örneklemesi sıralı bir şekilde uygulanabilir:

$n = 1, 2, \ldots$ için;

- $i = 1, \ldots, N$ için,
  - $n = 1$ ise $X_1^{(i)} \sim q_1(\cdot)$ üretilir, $w_1(X_1^{(i)}) = \frac{\pi_1(X_1^{(i)})}{q_1(X_1^{(i)})}$ hesaplanır.
  - $n \geq 2$ ise $X_n^{(i)} \sim q_n(\cdot|X_{1:n-1}^{(i)})$ üretilir, $X_{1:n}^{(i)} = (X_{1:n-1}^{(i)}, X_n^{(i)})$ oluşturulur, ve

$$w_n(X_{1:n}^{(i)}) = w_{n-1}(X_{1:n-1}^{(i)}) \frac{\widehat{\pi}_n(x_{1:n})}{\widehat{\pi}_{n-1}(x_{1:n-1}) q_n(X_n^{(i)}|X_{1:n-1}^{(i)})}.$$

- Öz-düzgeleyici önem ağırlıkları: $i = 1, \ldots, N$ için

$$W_n^{(i)} = \frac{w_n(X_{1:n}^{(i)})}{\sum_{i=1}^N w_n(X_{1:n}^{(i)})}.$$

SMM için sıralı önem örnekleyicisi

Hedef dağılımlar: $\pi_n(x_{1:n}) \propto \widehat{\pi}_n(x_{1:n}) = p(x_{1:n}, y_{1:n})$

$$p(x_{1:n}, y_{1:n}) = \eta(x_1) g(y_1|x_1) \prod_{t=2}^{n} f(x_t|x_{t-1}) g(y_t|x_t)$$

$p(x_{1:n}|y_{1:n})$ özyinelemeli olarak yazılabilir:

$$p(x_{1:n}, y_{1:n}) = p(x_{1:n-1}, y_{1:n-1}) f(x_n|x_{n-1}) g(y_n|x_n)$$

q_n sıralı olarak gözlemlere göre ayarlanabiliir. Örneğin,

$$\begin{aligned} q_n(x_{1:n}|y_{1:n}) &= q(x_1|y_1) \prod_{t=2}^{n} q(x_t|x_{t-1}, y_t) \\ &= q_{n-1}(x_{1:n-1}|y_{1:n-1}) q(x_n|x_{n-1}, y_n) \end{aligned}$$

Önem ağırlıkları:

$$w_n(x_{1:n}) = w_{n-1}(x_{1:n-1}) \frac{f(x_n|x_{n-1}) g(y_n|x_n)}{q(x_n|x_{n-1}, y_n)}.$$

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Parçacık süzgeci

Ağırlık bozulması sorunu

n arttıkça çok az sayıda $X_{1:n}^{(i)}$'nin önem ağırlıkları $w_n(X_{1:n}^{(i)})$ diğerlerininkine göre çok büyük olacaktır.

Dolayısıla, $W_n^{(i)}$ öz-düzgelenmiş ağırlıkarindan çok azı 1'e yakın olacak, diğerleri 0'a yaklaşacaktır.

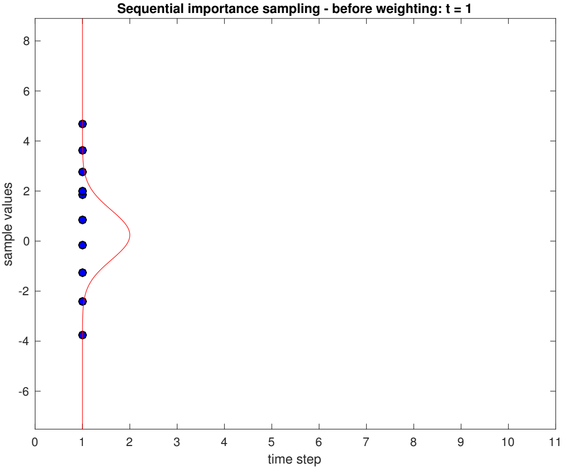
Limitte, $W_n^{(i)}$'lerden bir tanesi 1, diğerleri 0 olacaktır.

Bu soruna, ağırlık bozulması sorunu denir.

Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

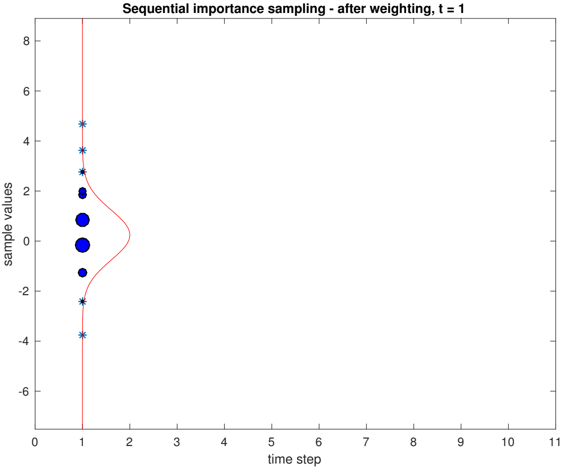
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

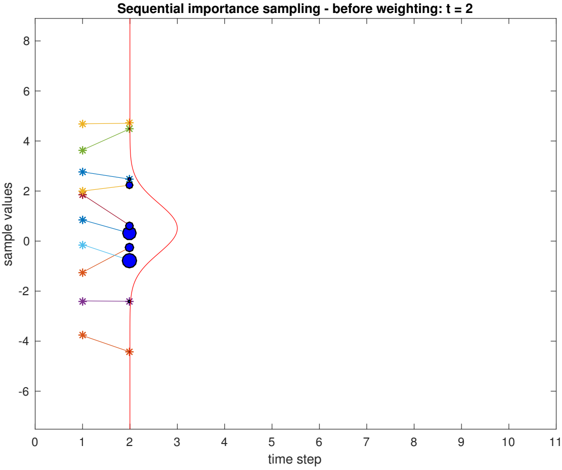
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

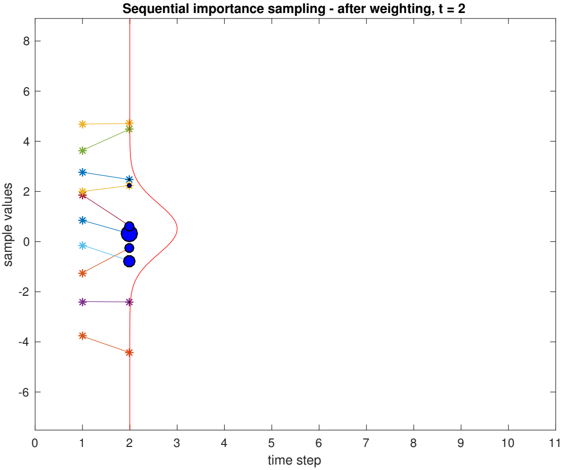
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

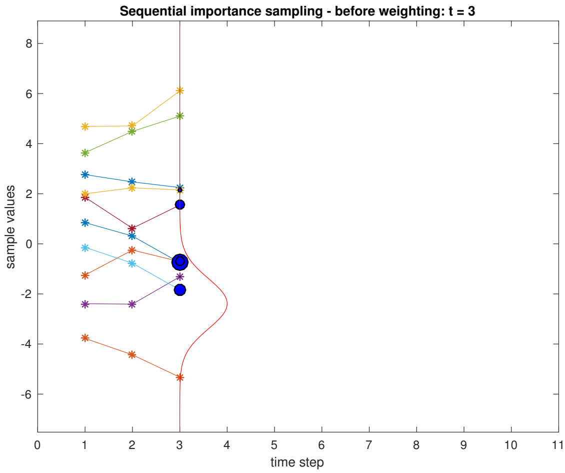
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

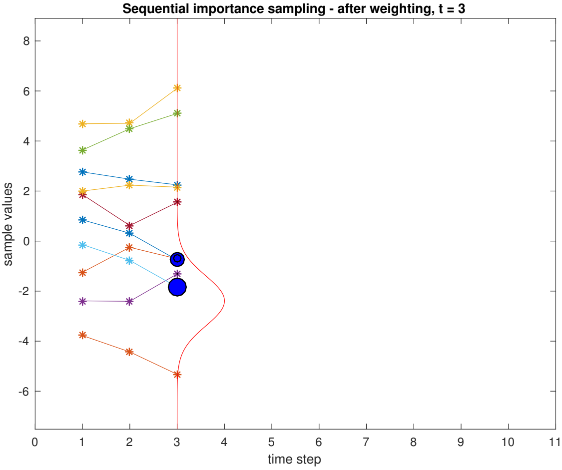
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

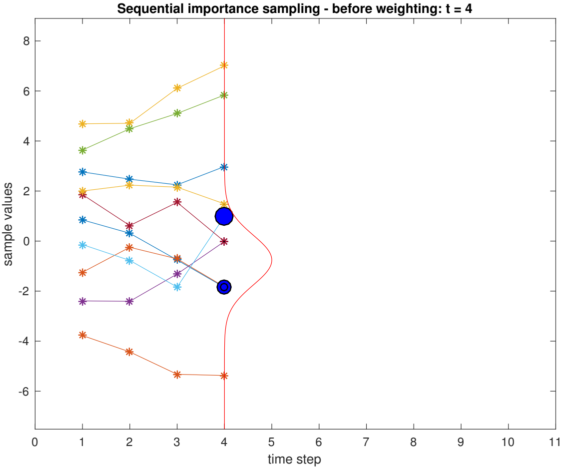
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

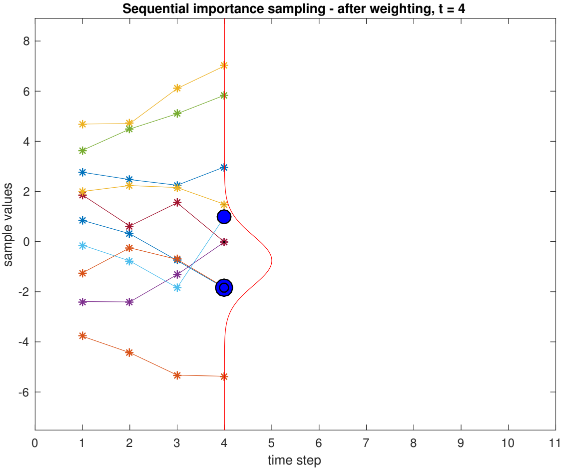
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

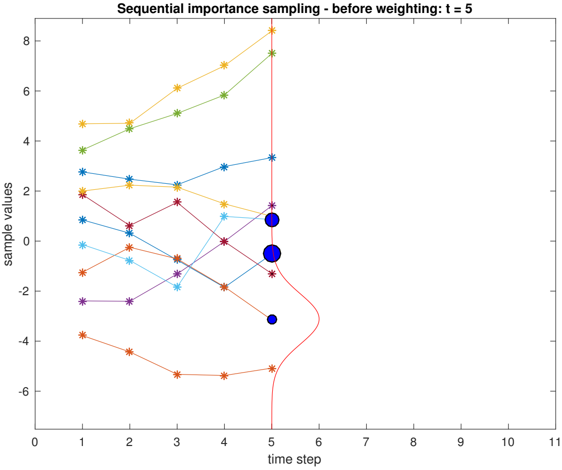
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

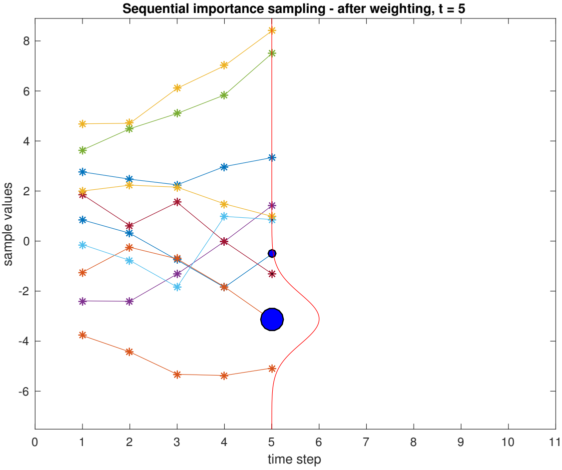
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

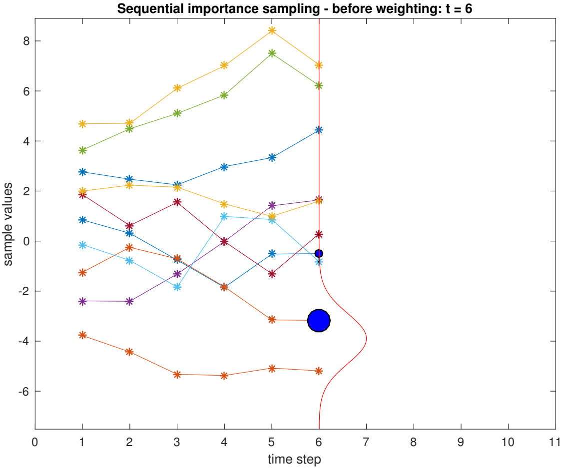
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

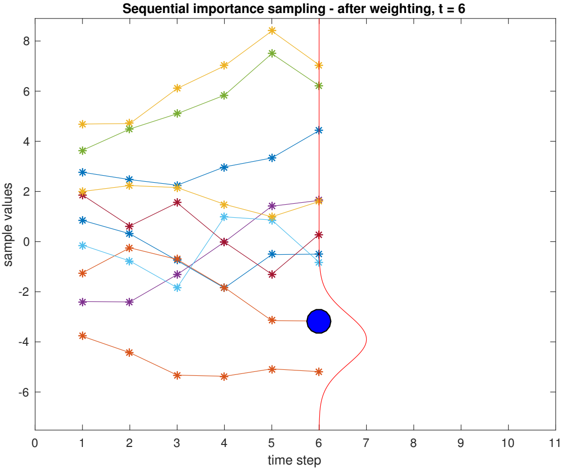
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

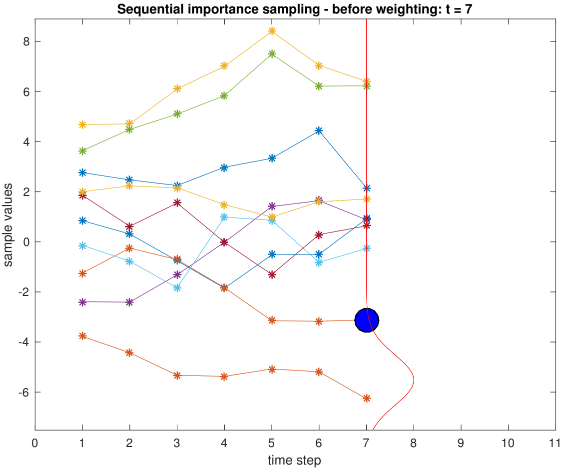
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

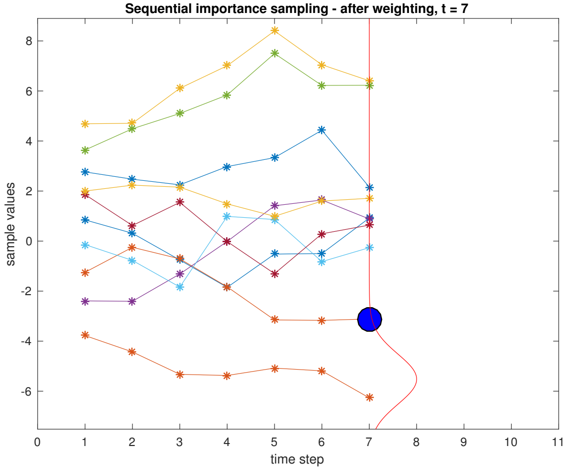
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

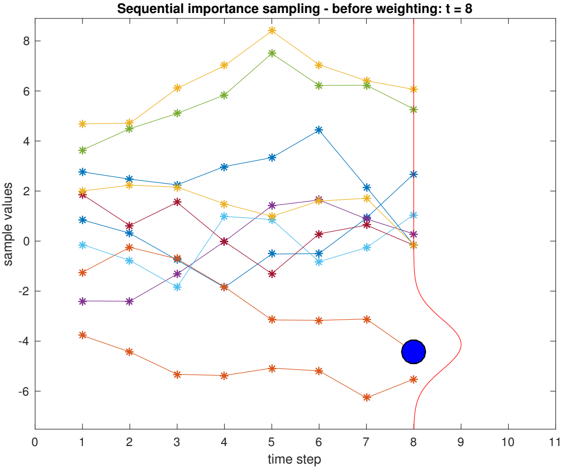
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

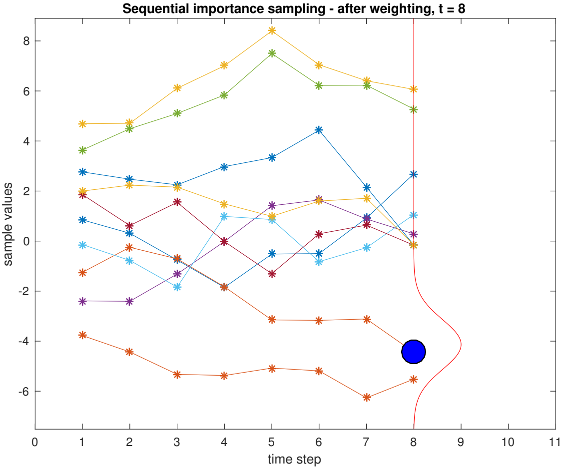
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

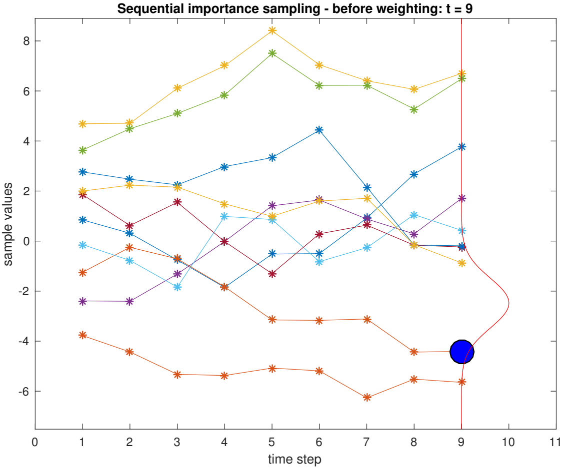
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

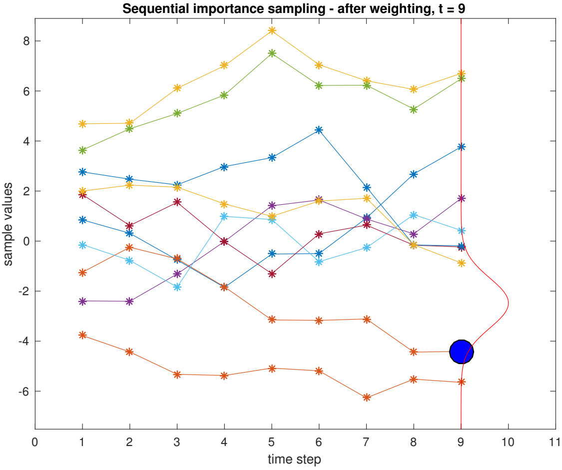
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

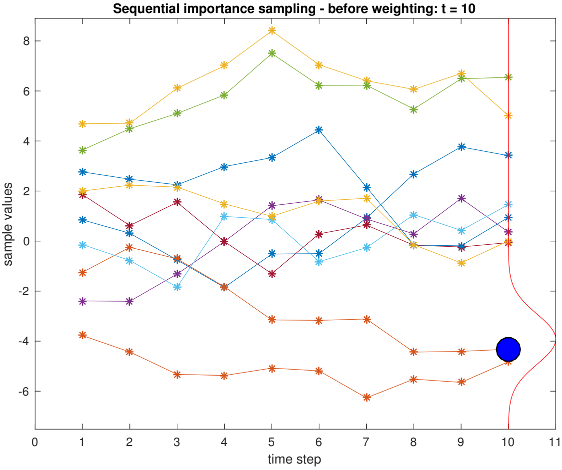
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

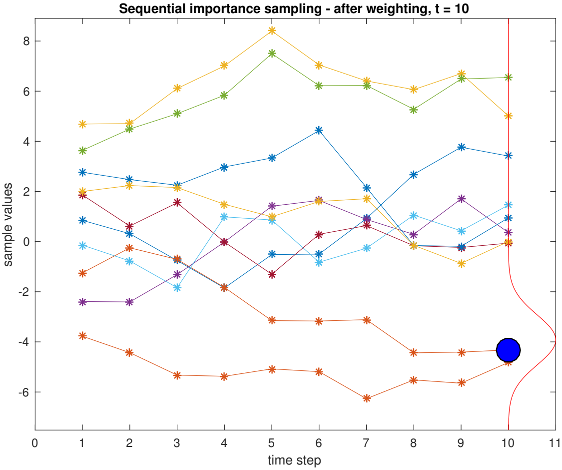
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Ağırlık bozulması sorunu: Örnek

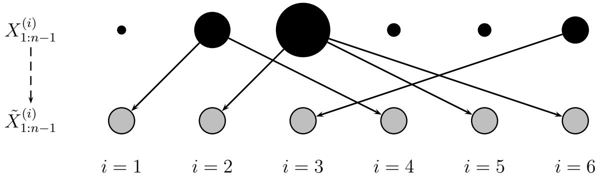
Doğrusal Gauss saklı Markov modeli

$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



Yeniden örnekleme → Parçacık süzgeci

Öz-düzgelenmiş ağırlıkları $W_{n-1}^{(1)}, \dots, W_{n-1}^{(N)}$ olan örnekler: $X_{1:n-1}^{(1)}, \dots, X_{1:n-1}^{(N)}$.



$$\mathbb{P}(\widetilde{X}_{1:n-1}^{(i)} = X_{1:n-1}^{(j)}) = W_{n-1}^{(j)}, \quad i, j = 1, \dots, N.$$

Artık yolumuza $1/N$ eşit ağırlıklı $\widetilde{X}_{1:n-1}^{(1)}, \dots, \widetilde{X}_{1:n-1}^{(N)}$ ile devam ediyoruz.

Parçacık süzgeci: Sıralı örnekleme yöntemine yeniden örnekleme adımının eklenmesiyle elde edilen yöntem.

SMM için parçacık süzgeci

Hedef dağılımlar: $\pi_n(x_{1:n}) \propto \widehat{\pi}_n(x_{1:n}) = p(x_{1:n}, y_{1:n})$

$$p(x_{1:n}, y_{1:n}) = \eta(x_1) g(y_1|x_1) \prod_{t=2}^{n} f(x_t|x_{t-1}) g(y_t|x_t)$$

Parçacık süzgeci:

$n = 1$ için;

$i = 1, \ldots, N$ için $X_1^{(i)} \sim q(\cdot|y_1)$ örneklenir ve $W_1^{(i)} \propto \frac{\eta(X_1^{(i)}) g(y_1|X_1^{(i)})}{q(X_1^{(i)}|y_1)}$ hesaplanır.

$n = 2, 3, \ldots$ için,

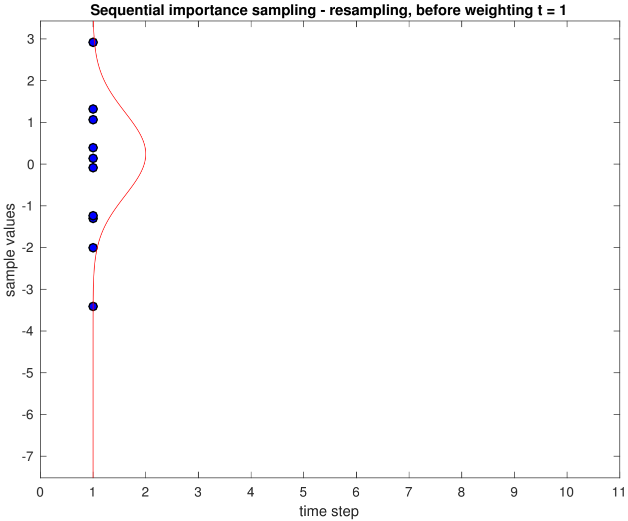
- Yeniden örnekleme ile $\widetilde{X}_{1:n-1}^{(1)}, \ldots, \widetilde{X}_{1:n-1}^{(N)}$ üretilir:

$$\mathbb{P}(\widetilde{X}_{1:n-1}^{(i)} = X_{1:n-1}^{(j)}) = W_{n-1}^{(j)}, \quad i, j = 1, \ldots, N.$$

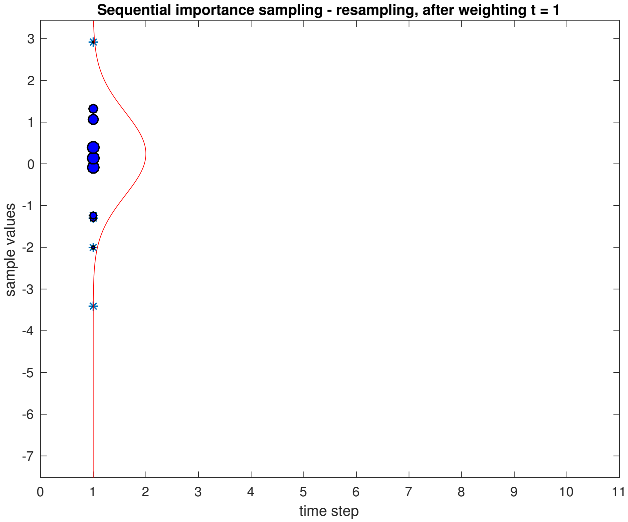
- $i = 1, \ldots, N$ için, $X_n^{(i)} \sim q_n(\cdot|\widetilde{X}_{n-1}^{(i)}, y_n)$ örneklenir, $X_{1:n}^{(i)} = (\widetilde{X}_{1:n-1}^{(i)}, X_n^{(i)})$ oluşturulur.
- Bu parçacıkların ağırlıkları

$$W_n^{(i)} \propto \frac{f(X_n^{(i)}|\widetilde{X}_{n-1}^{(i)}) g(y_n|X_n^{(i)})}{q(X_n^{(i)}|\widetilde{X}_{n-1}^{(i)}, y_n)}.$$

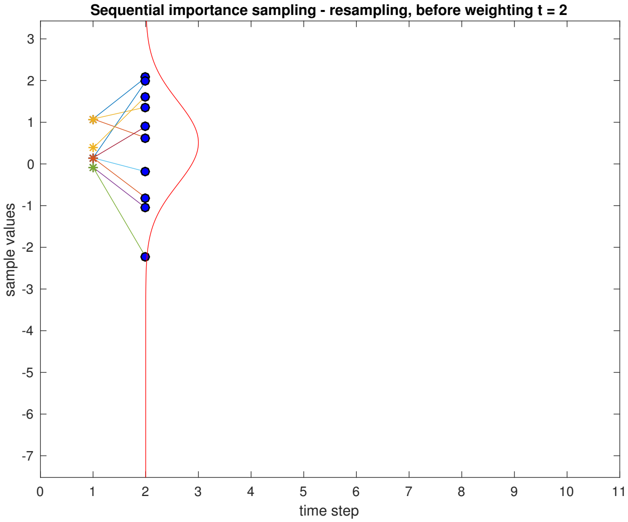
Parçacık süzgeci



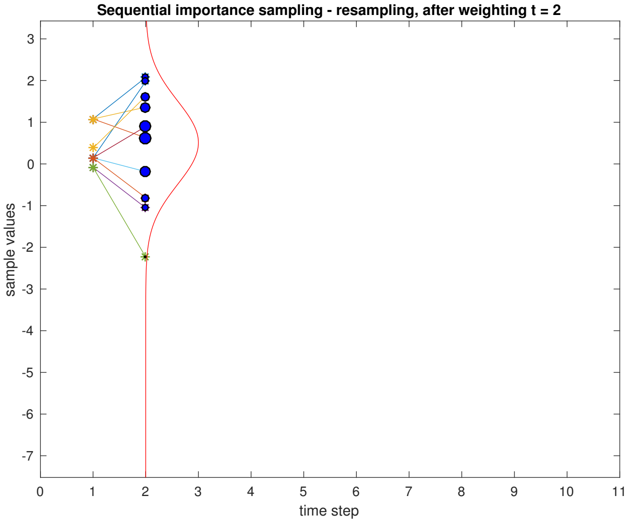
Parçacık süzgeci



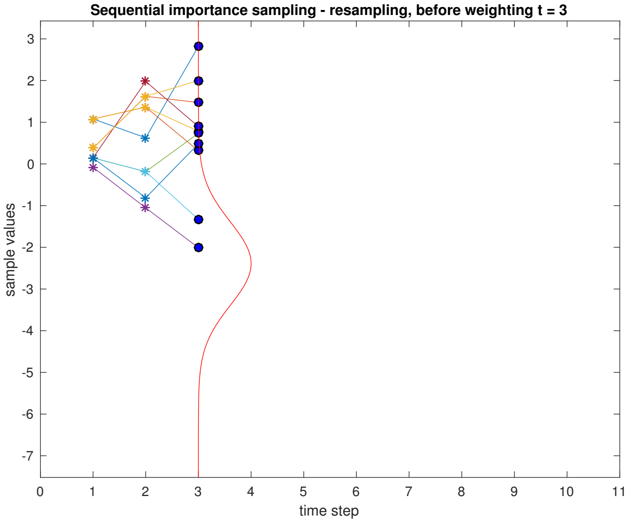
Parçacık süzgeci



Parçacık süzgeci

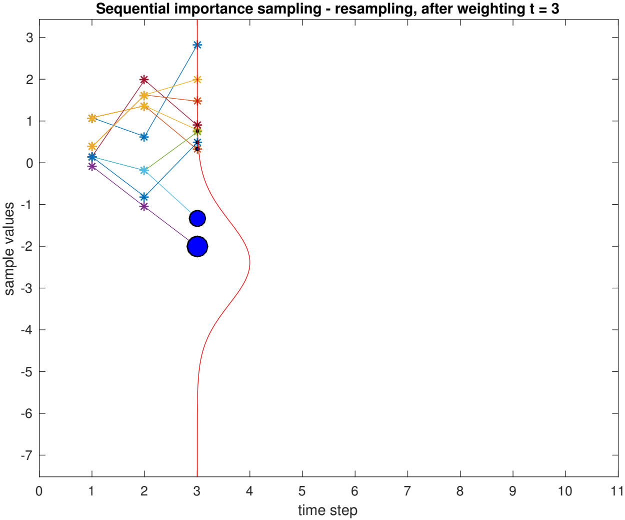


Parçacık süzgeci

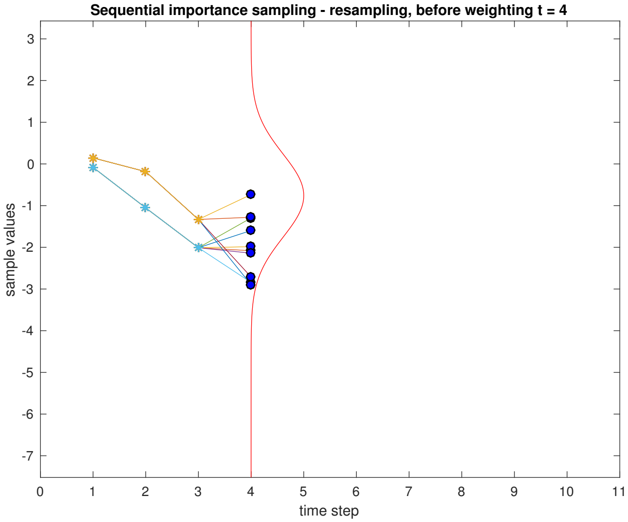


Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

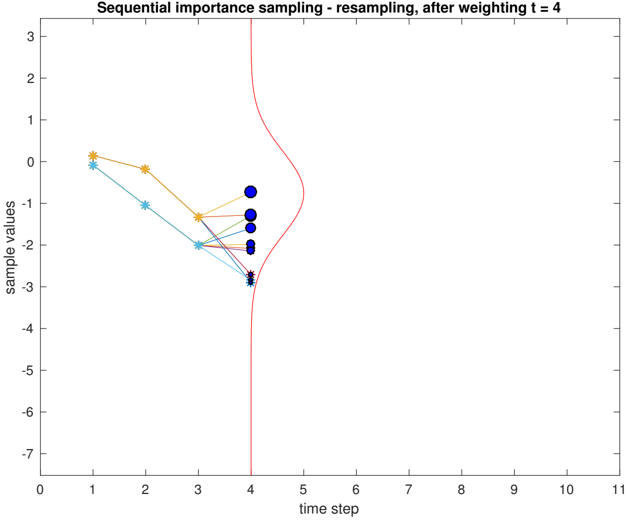
Parçacık süzgeci



Parçacık süzgeci

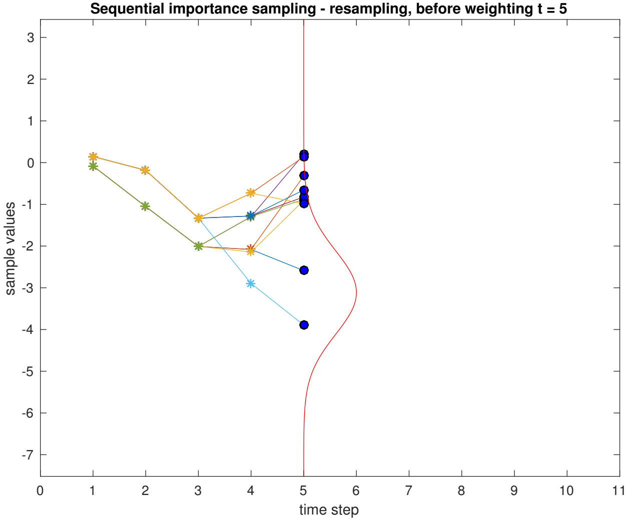


Parçacık süzgeci

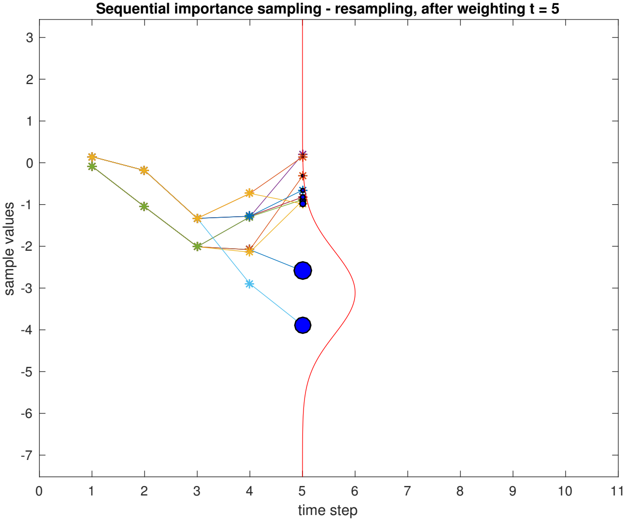


Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Parçacık süzgeci

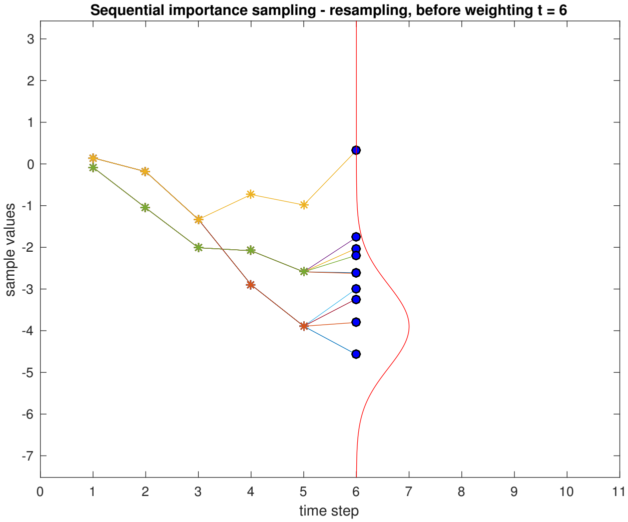


Parçacık süzgeci

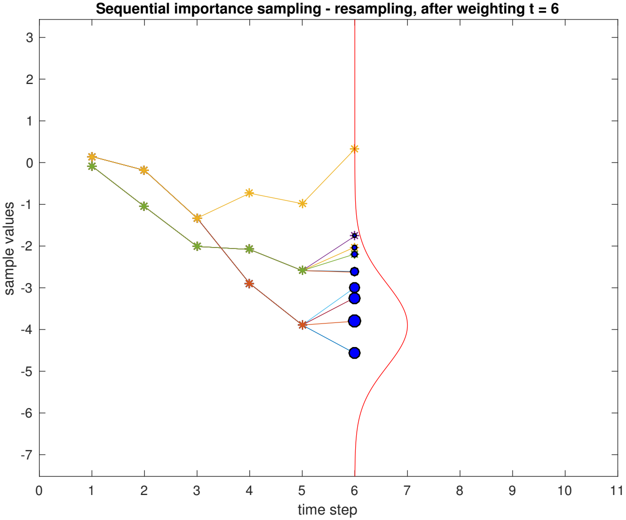


Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

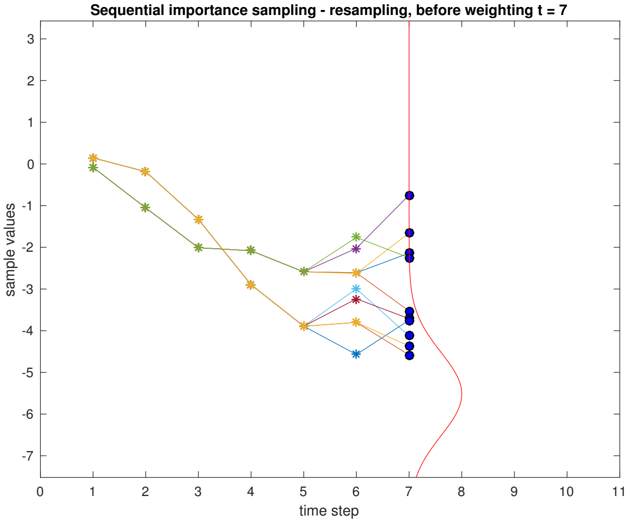
Parçacık süzgeci



Parçacık süzgeci

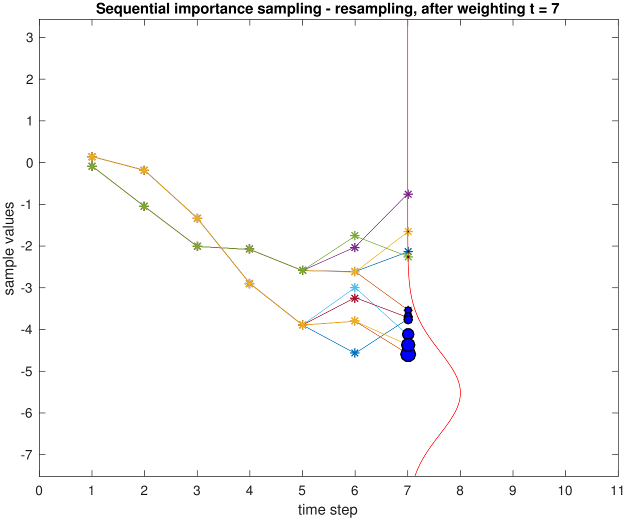


Parçacık süzgeci



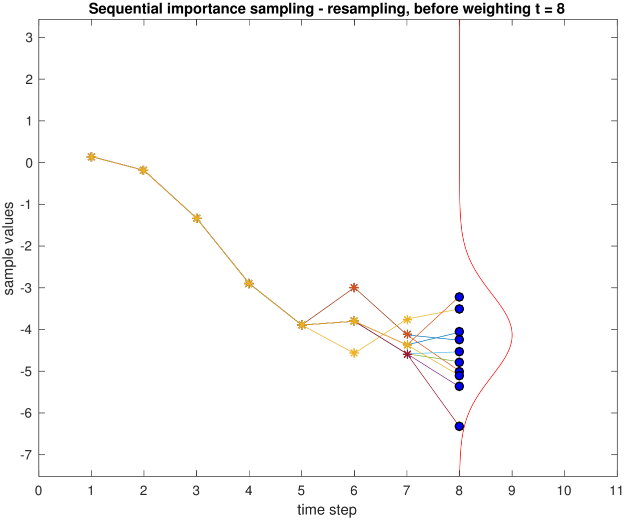
Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Parçacık süzgeci



Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

Parçacık süzgeci

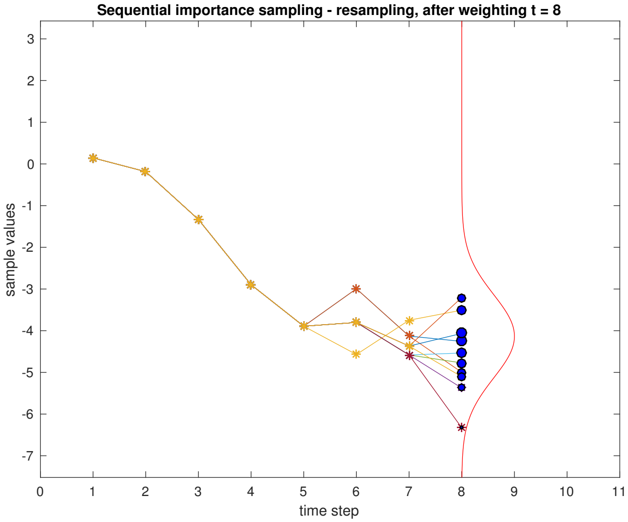


Monte Carlo

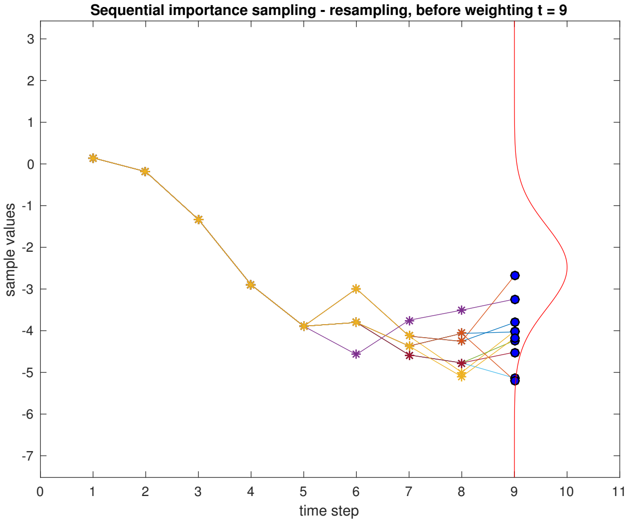
Sinan Yıldırım

Giriş
Örneklerin ortalaması
Monte Carlo
Kesin örnekleme yöntemleri
Tersini alma yöntemi
Dönüştürme yöntemi
Birleştirme yöntemi
Reddetme örneklemesi
Markov zinciri Monte Carlo
Gibbs örneklemesi
Metropolis-Hastings
Önem örneklemesi
Sıralı Monte Carlo
Sıralı önem örneklemesi
Parçacık süzgeci

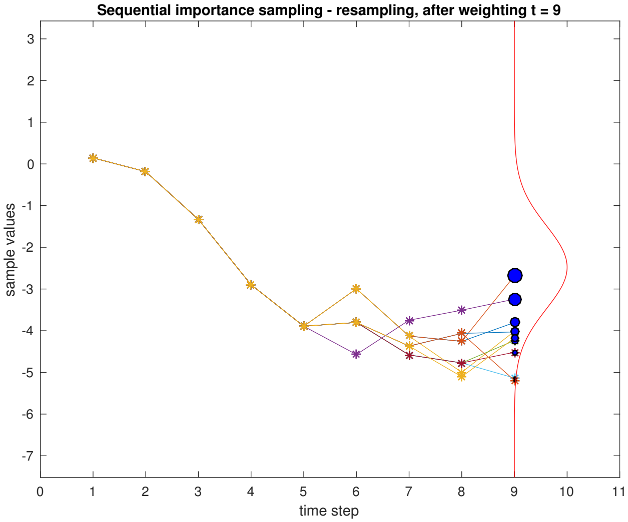
Parçacık süzgeci



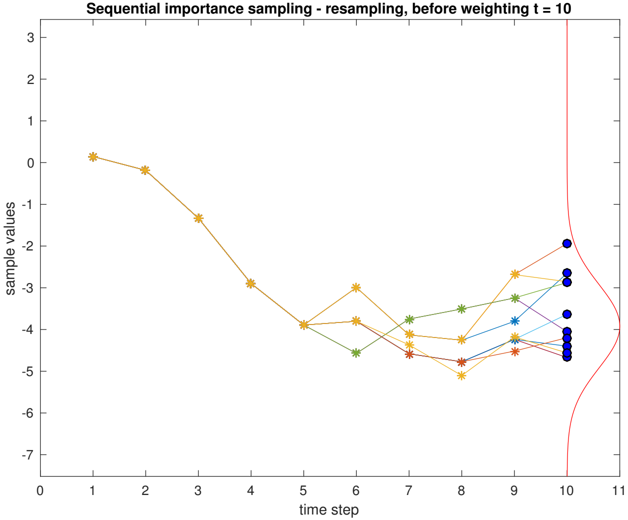
Parçacık süzgeci



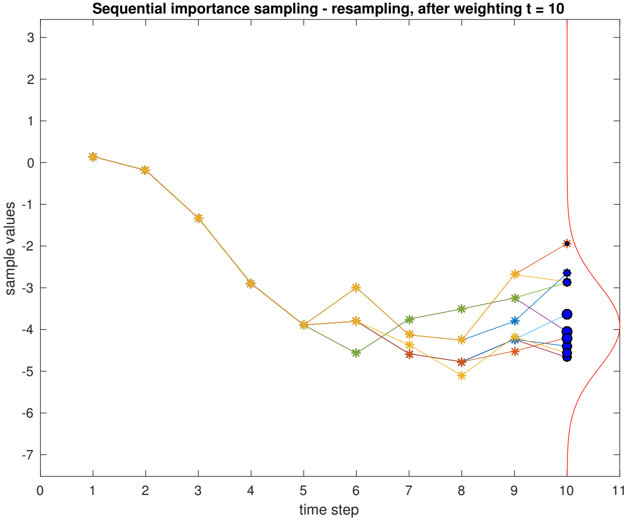
Parçacık süzgeci



Parçacık süzgeci



Parçacık süzgeci



Yeniden örnekleme: Yol bozulması sorunu

Ağırlık bozulması sorununu yeniden örnekleme ile giderilebilir.

Ancak yeniden örnekleme yol bozulması sorunu yaratır.

Art arda yeniden örneklemeler sebebiyle önceki zamanlara ait parçacık sayısı gitgide düşer.