

# Monte Carlo Yöntemleri

Sinan Yıldırım

MDBF, Sabancı Üniversitesi

May 9, 2017

## Giriş

- Buffon'un iğnesi

- Örneklerin ortalaması

- Monte Carlo

## Kesin örnekleme yöntemleri

- Tersini alma yöntemi

- Dönüştürme yöntemi

- Birleştirme yöntemi

- Reddetme örnekleme

## Önem örnekleme

## Markov zinciri Monte Carlo

- Metropolis-Hastings

- Gibbs örnekleme

## Sıralı Monte Carlo

- Sıralı önem örnekleme

- Parçacık süzgeci

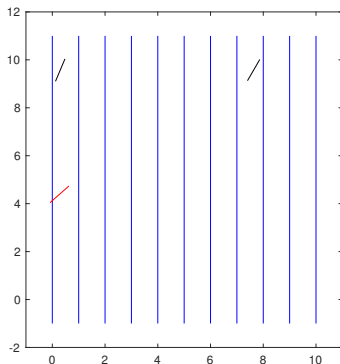
# Giriş

Buffon'un iğnesi

## Örnek: Buffon'un iğnesi

Eşit aralıklı (1 cm) paralel çizgileri olan bir masaya 1 cm'lik bir iğne rastgele atılıyor.

İğnenin masadaki çizgilerden birini kesmesi ihtimali nedir?



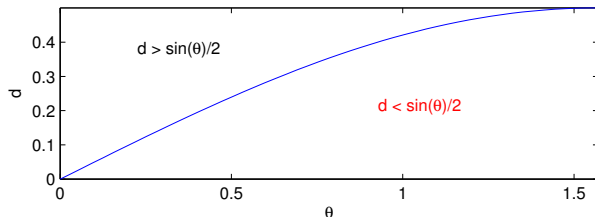
## Örnek: Buffon'un iğnesi

Bu olasılık kesin olarak hesaplanabilir:

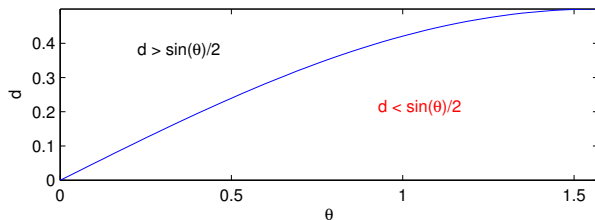
- ▶ İğnenin ortasının en yakın çizgiye uzaklığı  $d$  olsun.
- ▶ İğnenin çizgilerle yaptığı küçük açı  $\theta \in (0, \pi/2)$  olsun.

İğnenin çizgilerden birini kesmesi şu şartla olur:

$$\frac{d}{\sin \theta} < \frac{1}{2} \quad (1)$$



## Örnek: Buffon'un iğnesi



$d, \theta$  bağımsız ve  $d \in [0, 1/2]$  ve  $\theta \in [0, \pi/2]$  arasında homojen dağılır:

$$p(d, \theta) = \begin{cases} \frac{4}{\pi}, & (d, \theta) \in [0, 1/2] \times [0, \pi/2] \\ 0, & \text{else} \end{cases}$$

$A = \{(d, \theta) : d/\sin \theta < 1/2\} = \{(d, \theta) : d < \sin \theta/2\}$  kümesi, figürdeki eğrinin altındaki alana karşılık gelir.  $X = (d, \theta)$  dersek, bu kümenin olasılığı

$$\mathbb{P}(X \in A) = \int \int_A p(r, \theta) dr d\theta = \frac{2}{\pi}$$

## Örnek: Buffon'un iğnesi - Monte Carlo yaklaşımı

$\mathbb{P}(X \in A)$ 'ı yaklaşık hesaplamak için Monte Carlo deneyi:

$X^{(i)} = (d^{(i)}, \theta^{(i)})$ ,  $i = 1, \dots, N$  örnekleri üretilir:

$$d^{(i)} \sim \text{Unif}(0, 1/2), \quad \theta^{(i)} \sim \text{Unif}(0, \pi/2),$$

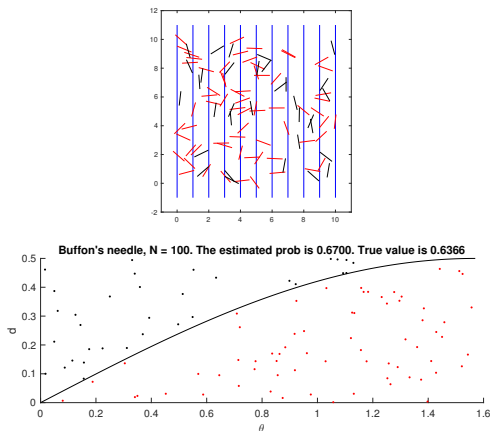
$\mathbb{P}(X \in A)$  olasılığı aşağıdaki gibi kestirilir:

$$\begin{aligned} \mathbb{P}(X \in A) &\approx \frac{1}{N} \sum_{i=1}^N \mathbb{I}_A(X^{(i)}) \\ &= \frac{1}{N} \sum_{i=1}^N \mathbb{I}(d^{(i)} < \sin(\theta^{(i)})/2) \end{aligned}$$



# Örnek: Buffon'un iğnesi - Monte Carlo yaklaşımı

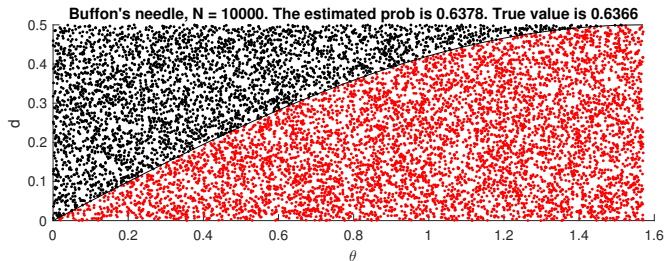
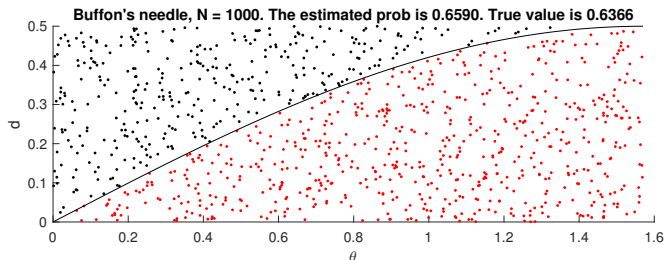
$N = 100$  atış ve karşılık geldikleri  $d, \theta$  değerleri.



$$\mathbb{P}(X \in A) \approx \frac{\text{number of red dots}}{\text{total number of dots}}.$$

# Örnek: Buffon'un iğnesi - Monte Carlo yaklaşımı

Büyük sayılar kanunu  $N$  arttıkça kestirimin  $2/\pi$ 'ye yaklaştığını öngörür.



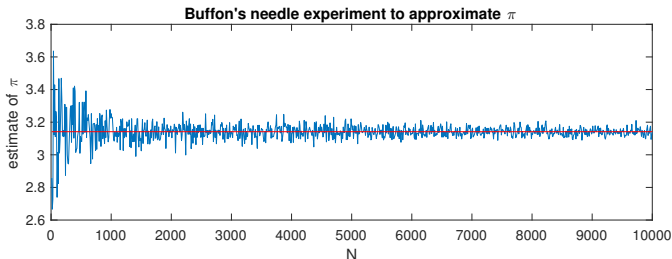
# $\pi$ 'nin kestirimi

$\pi$  ile  $\mathbb{P}(X \in A)$  arasındaki ilişki:

$$\pi = \frac{2}{\mathbb{P}(X \in A)},$$

$\pi$ 'nin Monte Carlo kestirimi:

$$\begin{aligned}\pi &\approx 2 \times \frac{N}{\sum_{i=1}^N \mathbb{I}(d^{(i)} < \sin(\theta^{(i)})/2)} \\ &= 2 \times \frac{\text{toplam nokta sayısı}}{\text{kırmızı nokta sayısı}}\end{aligned}$$



## 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:

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

Örneklerin *bağımsız ve özdeş dağılımlı* olduğunu ve  $\pi$  olasılık dağılımından geldiğini varsayalım:

$$X_1, \dots, X_N \stackrel{\text{i.i.d.}}{\sim} \pi.$$

Ayrıca  $\pi$  dağılımı bilinmiyor olsun.

## $\pi$ 'ye göre ortalama değer

$X$ 'in  $\pi$  dağılımına göre beklentisini  $X^{(1:N)}$  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  $X, x_1, x_2, \dots$  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  $\varphi : \mathcal{X} \rightarrow \mathbb{R}$  fonksiyonunun  $\pi$ '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:  $\mathbb{I}_A : \mathcal{X} \rightarrow \{0, 1\}$

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

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 yaklaşılabılır.



# Monte Carlo

# Monte Carlo: Ana fikir

Şimdi şu senaryoyu düşünelim:  $\pi$  dağılımını biliyoruz, ama  $\pi$ 'den gelen örneklerimiz yok.

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

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

Eğer  $X^{(1)}, \dots, X^{(N)} \sim \pi$  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,  $\pi_{\text{MC}}^N(\varphi)$ 'nin 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

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

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

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

# Monte Carlo'nun gerekçelendirilmesi - Merkezi limit teoremi

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

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

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

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

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

# 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.

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

# İleri yöntemlere ihtiyaç

Çoğu problemde, tek sorun integralin alınamazlığı değil.

- $\pi'$ 'den örnekleme yapmak homojen dağılım kadar kolay olmayabilir.

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

- $\pi'$ 'den örnekleme yapmak imkansız olabilir.

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

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

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

# Kesin örnekleme yöntemleri



# 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  $\text{Unif}(0, 1)$  dağılımından gelen sayılar olsun. Bu sayıları kullanarak herhangi bir  $\pi$  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.

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

# Tersini alma yöntemi

$F$ 'nin genelleştirilmiş tersi:

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

$X$  ayrık ise ve  $x_1, x_2, \dots$  değerlerini alıyorsa:

$$G(u) = x_{i^*}, \quad i^* = \min\{i : F(x_i) \geq u\} \Leftrightarrow F(x_{i^*-1}) < u \leq F(x_{i^*})$$

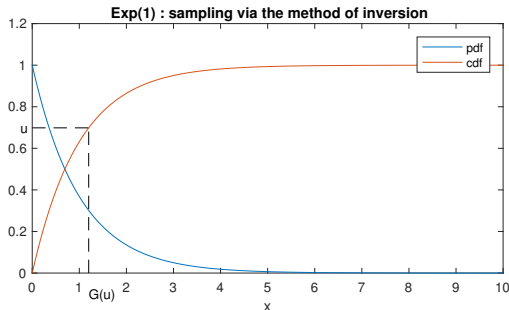
$X$  sürekli ise ve  $\pi(x) > 0$  şeklinde bir olasılık yoğunluk fonksiyonu varsa ( $F$ 'te sıçrama ve düz alanlar yok),  $F$  monoton artandır ve tersi  $G = F^{-1}$  alınabilir.

$$G(u) = F^{-1}(u)$$

## Örnek: Üssel dağılım

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

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

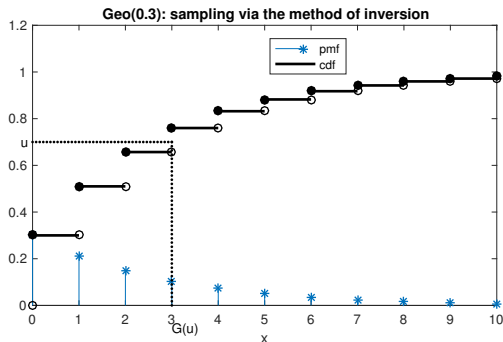


O halde,  $\text{Exp}(\lambda)$ 'dan  $U \sim \text{Unif}(0, 1)$  ve  $X = -\log(1 - U)/\lambda$  şeklinde örnek üretebiliriz.

# 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  $X = \left\lceil \frac{\log(1-U)}{\log(1-\rho)} \right\rceil$  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,  $U \sim \text{Unif}(0, 1)$  üretilip  $U$ 'yu

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

şeklinde dönüştürebiliriz.



# Dönüştürme yöntemi: genel

Elimizde olasılık yoğunluk fonksiyonu  $p_X(x)$  olan  $m$  boyutlu  $X \in \mathcal{X} \subseteq \mathbb{R}^m$  rassal değişkeni olsun.

Tersi alınabilir bir  $g : \mathcal{X} \rightarrow \mathcal{Y} \subseteq \mathbb{R}^m$  kullanarak  $X$ 'i dönüştürelim:

$$Y = (Y_1, \dots, Y_m) = g(X_1, \dots, X_m)$$

Soru:  $Y$ 'nin olasılık yoğunluk dağılımı  $p_Y(y)$  ne olur?

$y = (y_1, \dots, y_m) = g(x_1, \dots, x_m)$  kullanarak, Jakobian'ı tanımlayalım

$$J(y) = \det \frac{\partial g^{-1}(y)}{\partial y} = \det \frac{\partial x}{\partial y} = \det \frac{\partial(x_1, \dots, x_m)}{\partial(y_1, \dots, y_m)} = \det \begin{bmatrix} \partial x_1 / \partial y_1 & \dots & \partial x_1 / \partial y_m \\ \vdots & \ddots & \vdots \\ \partial x_m / \partial y_1 & \dots & \partial x_m / \partial y_m \end{bmatrix}$$

$Y$ 'nin olasılık yoğunluk fonksiyonu:

$$p_Y(y) := p_X(g^{-1}(y)) |J(y)|$$

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

Standart Gauss dağılımı (normal dağılım)  $\mathcal{N}(0, 1)$ 'in 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

İlk önce

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

üretilir, sonra

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

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

## Box-Muller yöntemi: Kanıt

Bu yöntem doğrudan değişkenlerin dönüştürülmesine dayanır:

$$(R, \Theta) = (X_1^2 + X_2^2, \arctan(X_2/X_1))$$

Jakobian'ın  $(r, \theta) = (x_1^2 + x_2^2, \arctan(x_2/x_1))$ 'deki değeri

$$J(r, \theta) = \begin{vmatrix} \partial r / \partial x_1 & \partial r / \partial x_2 \\ \partial \theta / \partial x_1 & \partial \theta / \partial x_2 \end{vmatrix} = \begin{vmatrix} 2x_1 & 2x_2 \\ \frac{1}{1+(y_2/y_1)^2} \frac{-y_2}{y_1^2} & \frac{1}{1+(y_2/y_1)^2} \frac{1}{y_1} \end{vmatrix} = 2$$

$(X_1, X_2)$ 'nin olasılık yoğunluk fonksiyonu, formülü kullanarak,

$$\begin{aligned} p_{X_1, X_2}(x_1, x_2) &= p_R(r) p_\Theta(\theta) |J(r, \theta)| \\ &= p_R(x_1^2 + x_2^2) p_\Theta(\arctan(x_2/x_1)) |J(r, \theta)| \\ &= \frac{1}{2} e^{-\frac{1}{2}(x_1^2 + x_2^2)} \frac{1}{2\pi} 2 \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} \\ &= \phi(x_1; 0, 1) \phi(x_2; 0, 1) \end{aligned}$$

olarak bulunabilir, ki bu da iki Gauss dağılımının çarpımı olduğu için  $X_1, X_2 \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$  olduğu görülür.

# Çokdeğişkenli Gauss dağılımı

$n \times 1$  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,  $|\cdot|$  determinanti simgeler.

# Çokdeğişkenli Gauss dağılımı

Varsayalım ki  $X = (X_1, \dots, X_n)^T \sim \mathcal{N}(\mu, \Sigma)$  olsun.

Kertesini  $m \leq n$  olan  $m \times n$  bir matris ve bir  $m \times 1$   $\eta$  vektörünü kullanarak  $X$ 'i

$$Y = AX + \eta$$

şeklinde dönüştürelim.

$Y$ ,  $X$ 'in doğrusal dönüştürülmüş hali olduğu için  $Y$  de Gauss dağılımına sahiptir.

$$\begin{aligned}\mathbb{E}(Y) &= \mathbb{E}(AX + \eta) \\ &= A\mathbb{E}(X) + \eta \\ &= A\mu + \eta\end{aligned}$$

$$\begin{aligned}\text{Cov}(Y) &= \mathbb{E}([Y - \mathbb{E}(Y)][Y - \mathbb{E}(Y)]^T) \\ &= \mathbb{E}([AX + \eta - (A\mu + \eta)][AX + \eta - (A\mu + \eta)]^T) \\ &= \mathbb{E}(A(X - \mu)(X - \mu)^T A^T) = A\text{Cov}(X)A^T \\ &= A\Sigma A^T\end{aligned}$$

Sonuç olarak,  $Y \sim \mathcal{N}(A\mu + \eta, A\Sigma A^T)$  yazabiliriz.

# Çokdeğişkenli Gauss dağılımı: örnekleme

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

Önce  $R_1, \dots, R_n \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1)$  üretilir böylece

$$R = (R_1, \dots, R_n) \sim \mathcal{N}(0_n, I_n)$$

sağlanmış olur.

Sonra, Cholesky ayrıştırması kullanılarak

$$\Sigma = AA^T$$

eşitliğini sağlayan  $A$  matrisi bulunur.

Son olarak

$$X = AR + \mu$$

değişkeni üretilir.

## 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,  $X \sim P$ '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?

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

Doğrudan  $P$ 'den örnekleme çok zor olabilir, ancak  $\Pi$  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  $\pi$ 'den gelir.

# Örnek: Karışım Gauss dağılımı

$K$  bileşeni olan, bileşenlerinin

- ▶ ortalama değerleri ve varyansları:  $(\mu_1, \sigma_1^2), \dots, (\mu_K, \sigma_K^2)$
- ▶ karışımındaki olasılık ağırlıkları  $w_1, \dots, w_K$  ( $w_1 + \dots + w_K = 1$  olacak şekilde)

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.

## Örnek: Mahremiyet gözetilen veri paylaşımı

Bir şirket  $D$  olan aylık talep miktarını mahremiyet sebebiyle gürültü olarak  $X$  şeklinde paylaşıyor.

Paylaşılan  $X$ 'in dağılımı

$$\pi(x) = \sum_d \left[ \frac{e^{-\lambda} \lambda^d}{d!} \right] \left[ \frac{1}{2b} \exp \left( -\frac{|x - d|}{b} \right) \right]$$

Bu paylaşım sürecinin Monte Carlo ile benzetimini yapmak istiyoruz.  $X \sim \pi$  nasıl üretilir?

Toplamdaki ilk terim  $\mathcal{PO}(\lambda)$ 'nin olasılık kütle fonksiyonunun  $d$ 'deki değeri, diğeri de  $\text{Laplace}(d, b)$ 'nin olasılık dağılım fonksiyonunun  $x$ 'teki değeri.

1.  $D \sim \mathcal{PO}(\lambda)$  üretilir,
2.  $X \sim \text{Laplace}(D, b)$  üretilir (veya  $V \sim \text{Laplace}(0, b)$  ve  $X = D + V$ ).
3.  $D$  atılır  $X$  tutulur.

## Reddetme örneklemeesi

# 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 örnekleme: 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

$$\begin{aligned}\mathbb{P}(\text{Kabul}) &= \int \mathbb{P}(\text{Kabul} | X' = x) p_{X'}(x) dx \\ &= \int \frac{\pi(x)}{Mq(x)} q(x) dx \\ &= \frac{1}{M} \int \pi(x) dx \\ &= \frac{1}{M},\end{aligned}\tag{2}$$

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

# Reddetme örnekleme: Neden çalışır?

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.

Yöntemin doğruluğunu göstermek için, üretilen  $X$ 'in dağılımının  $\pi$  olduğunu göstermek gerekir.

Bayes' teoremini kullanarak,

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

## Örnek: Gamma dağılımı

Örneklenecek dağılım:  $X \sim \Gamma(\alpha, 1)$ ,  $\alpha > 1$ . Olasılık yoğunluk fonksiyonu:

$$\pi(x) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)}, \quad x > 0.$$

Aracı dağılım olarak  $q_\lambda = \text{Exp}(\lambda)$ ,  $0 < \lambda < 1$ , seçelim.

$$q_\lambda(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

Bütün  $x \in \mathcal{X}$  için  $\pi(x) \leq M q_\lambda(x)$ 'i sağlayacak  $M$  bulunmalı:

$$\frac{\pi(x)}{q_\lambda(x)} = \frac{x^{\alpha-1} e^{(\lambda-1)x}}{\lambda \Gamma(\alpha)}$$

$x = (\alpha - 1)/(1 - \lambda)$ 'de enbüyütülür, dolayısıyla

$$M_\lambda = \frac{\left(\frac{\alpha-1}{1-\lambda}\right)^{\alpha-1} e^{-(\alpha-1)}}{\lambda \Gamma(\alpha)}$$

alınabilir. Kabul olasılığı

$$\frac{\pi(x)}{q_\lambda(x) M_\lambda} = \left(\frac{x(1-\lambda)}{\alpha-1}\right)^{\alpha-1} e^{(\lambda-1)x + \alpha - 1}$$



## Örnek: Gamma dağılımı

Kabul olasılığı  $1/M_\lambda$ ,  $M_\lambda$ 'yı önceden bulmuştuk:

$$M_\lambda = \frac{\left(\frac{\alpha-1}{1-\lambda}\right)^{\alpha-1} e^{-(\alpha-1)}}{\lambda \Gamma(\alpha)}$$

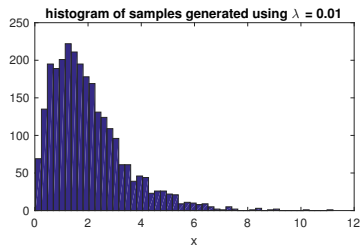
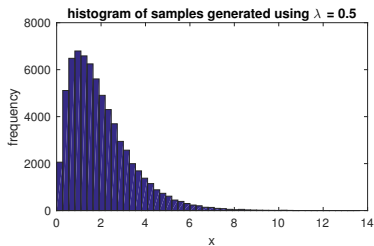
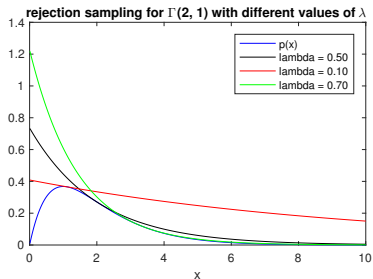
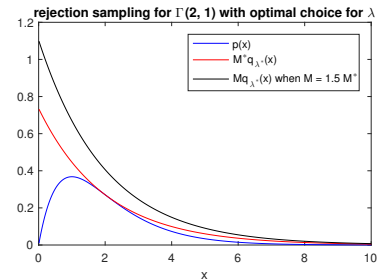
Şimdi de  $M_\lambda$ 'yı enküçültecek  $\lambda$ 'yı seçelim.  $M_\lambda$ ,  $\lambda^* = 1/\alpha$ 'da enküçültülür ve

$$M^* = \frac{\alpha^\alpha e^{-(\alpha-1)}}{\Gamma(\alpha)}.$$

elde edilir. Sonuç olarak  $\Gamma(\alpha, \beta)$ 'dan örnekleme yapmak için,

1.  $X' \sim \text{Exp}(1/\alpha)$  ve  $U \sim \text{Unif}(0, 1)$  örneklenir.
2. If  $U \leq (x/\alpha)^{\alpha-1} e^{(1/\alpha-1)x+\alpha-1}$ , ise  $X = X'$  alınır, değilse 1'e gidilir.

# Örnek: Gamma dağılımı



## $\pi(x)$ tam olarak bilinmediğinde

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

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

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

1.  $X' \sim q(\cdot)$  ve  $U \sim \text{Unif}(0, 1)$  üretilir.
2.  $U \leq \frac{\hat{\pi}(X')}{Mq(X')}$  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

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

Çarpımsal (ve çoğunlukla hesaplanamayan) sabit:

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

# Bilinmeyen sabite örnek: Bayesci çıkarım

$X'$ 'in  $Y = y$  verildiğindeki sonsal dağılımı

$$p_{X|Y}(x|y) \propto p_X(x)p_{Y|X}(y|x)$$

Çarpımsal (ve çoğunlukla hesaplanamayan) sabit:

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

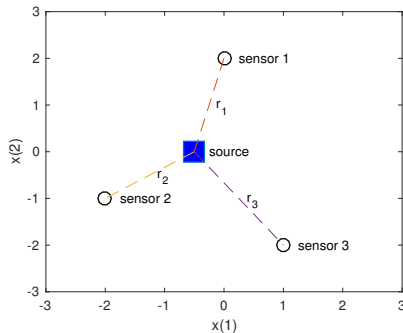
$\pi(x) = p_{X|Y}(x|y)$ 'dan örnekleme için reddetme örnekleme kullanılabilir.

Örnek: Eğer bütün  $x \in \mathcal{X}$  için  $p_{Y|X}(y|x) \leq M$ 'i sağlayacak bir  $M$  varsa, reddetme örnekleme bu  $M$  ile ve  $q(x) = p_X(x)$  ile kullanılabilir:

1.  $X' \sim q(\cdot)$  ve  $U \sim \text{Unif}(0, 1)$  üretilir,
2. Eğer  $U \leq p_{Y|X}(y|X')/M$  ise  $X = X'$  alınır; değilse 1.'e gidilir.

# 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.

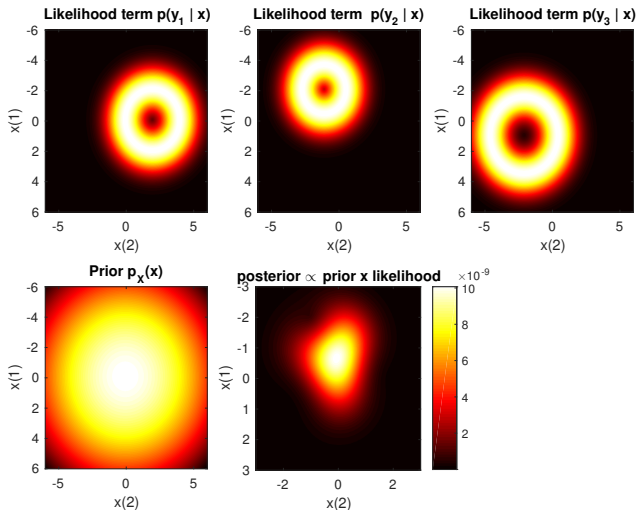
$$\pi(x) := p_{X|Y}(x|y) \propto \underbrace{p_X(x)p_{Y|X}(y|x)}_{\hat{\pi}(x)} = \underbrace{\phi(x; 0_2, \sigma_x^2 I_2)}_{p_X(x)} \underbrace{\prod_{i=1}^3 \phi(y_i; r_i, \sigma_y^2)}_{p_{Y|X}(y|x)}$$

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

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

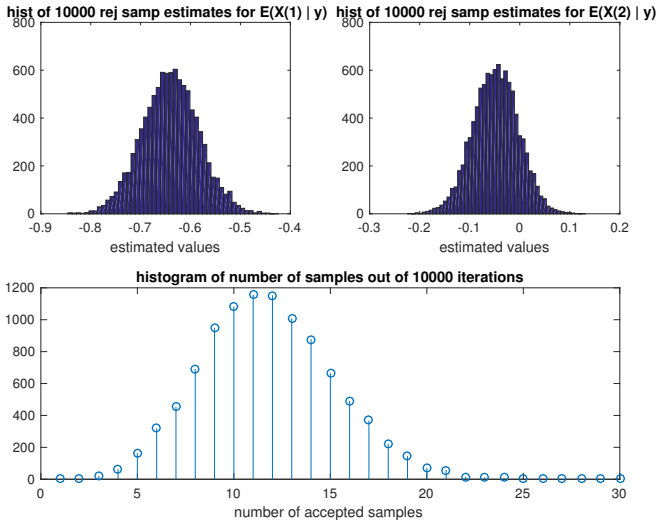
O halde,  $M = \frac{1}{(2\pi\sigma_y^2)^{3/2}}$  seçilmelidir.

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





# Hedef yer saptaması - reddetme örneklemesi $\sigma_x^2 = 100$ , $\sigma_y^2 = 1$



# Önem örneklemesi

# Önem örnekleme: 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

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

için  $\pi$ 'den örnekleme yapmamız gerekiyor.

Bir çok durumda  $X \sim \pi$  örnekleme 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  $w : \mathcal{X} \rightarrow \mathbb{R}$

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

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 örnekleme

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

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

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

# Önem örnekleme: Örnek

$(X, Y) \in \mathcal{X} \times \mathcal{Y}$  değişkenlerinin ortak dağılımının yoğunluk fonksiyonu  $p_{X,Y}(x, y)$  olsun

$$p_{X,Y}(x, y) = p_X(x)p_{Y|X}(y|x)$$

Bazı durumlarda  $p_Y(y)$  hesaplanmak istenebilir:

$$p_Y(y) = \int_{\mathcal{X}} p_{X,Y}(x, y) dx = \int_{\mathcal{X}} p_X(x)p_{Y|X}(y|x) dx = \mathbb{E}_{p_X}(p_{Y|X}(y|X))$$

Standard Monte Carlo kestirimi:

$$p_Y(y) \approx \frac{1}{N} \sum_{i=1}^N p_{Y|X}(y|X^{(i)}), \quad X^{(1)}, \dots, X^{(N)} \sim p_X(x).$$

Önem örnekleme ile kestirim:

$$p_Y(y) \approx \frac{1}{N} \sum_{i=1}^N \frac{p_X(X^{(i)})}{q(X^{(i)})} p_{Y|X}(y|X^{(i)}), \quad X^{(1)}, \dots, X^{(N)} \sim q(x).$$

# Önem örnekleme: en iyi $q(x)$

$q(x)$ 'yu seçme özgürlüğümüz var, o halde en iyileştirmeye çalışalım.

$$\mathbb{V}_q \left[ \pi_{\text{IS}}^N(\varphi) \right] = \frac{1}{N} \mathbb{V}_q [w(X)\varphi(X)]$$

Varyansı en küçük yapacak  $q(x)$

$$q(x) = \pi(x) \frac{|\varphi(x)|}{\pi(|\varphi|)}$$

Bu tercih ile elde edilen varyans:

$$\min_q \mathbb{V}_q \left[ \pi_{\text{IS}}^N(\varphi) \right] = \frac{1}{N} ([\pi(|\varphi|)]^2 - [\pi(\varphi)]^2) .$$

# Öz-düzgeleyici önem örnekleme

Önem örnekleme  $\pi(x) = \frac{\hat{\pi}(x)}{Z_p}$  olduğunda ve sadece  $\hat{\pi}(x)$  bilindiğinde yine uygulanabilir.

Önem fonksiyonu

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

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

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

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 örnekleme

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

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

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

Öz-düzgeleyici önem ağırlıkları:

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

Öz-düzgeleyici önem örnekleme

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

# Öz-düzgeleyici önem örnekleme: Bayesci çıkarım

Sonsal dağılım:

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

Hesaplanmak istenen beklenti değeri

$$\mathbb{E}(\varphi(X)|Y=y) = \int p_{X|Y}(x|y)\varphi(x)dx.$$

Öz-düzgeleyici önem örnekleme:

1.  $i = 1, \dots, N$  için;  $X^{(i)} \sim q(\cdot)$  örneklenir ve önem ağırlıkları hesaplanır

$$w(X^{(i)}) = \frac{p_X(X^{(i)})p_{Y|X}(y|X^{(i)})}{q(X^{(i)})}.$$

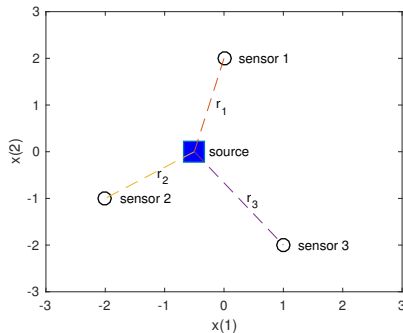
2.  $i = 1, \dots, N$  için;  $W^{(i)} = \frac{w(X^{(i)})}{\sum_{j=1}^N w(X^{(j)})}$  hesaplanır.

3.  $\mathbb{E}(\varphi(X)|Y=y) \approx \sum_{i=1}^N W^{(i)}\varphi(X^{(i)})$  hesaplanır.

Eğer  $q(x) = p_X(x)$  seçilirse, önem fonksiyonu  $w(x) = p_{Y|X}(y|x)$  olur.

# 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:

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

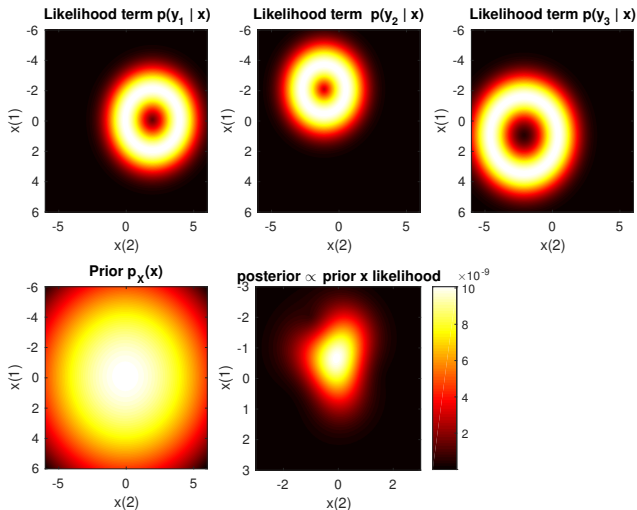
Amaç:  $Y = y = (y_1, y_2, y_3)$  verildiğinde  $p_{X|Y}(x|y)$ 'i bulmak ve  $\mathbb{E}(X|Y = y)$ 'yi hesaplamak.

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

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

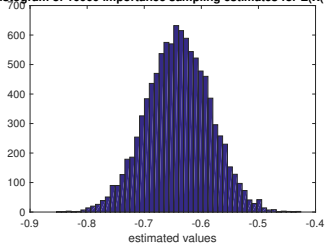
$$w(x) = \frac{\hat{\pi}(x)}{q(x)} = \frac{p_X(x)p_{Y|X}(y|x)}{p_X(x)} = p_{Y|X}(y|x) = \prod_{i=1}^3 \phi(y_i; r_i, \sigma_y^2)$$

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

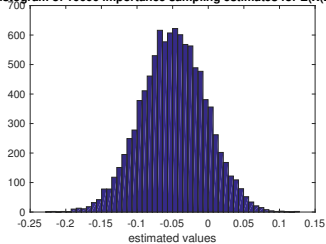


# Hedef yer saptaması

histogram of 10000 importance sampling estimates for  $E(X(1) | y)$



histogram of 10000 importance sampling estimates for  $E(X(2) | y)$



# Öz-düzgeleyici önem örnekleme: $q(x)$ 'in seçiminin önemi

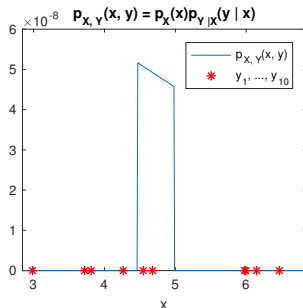
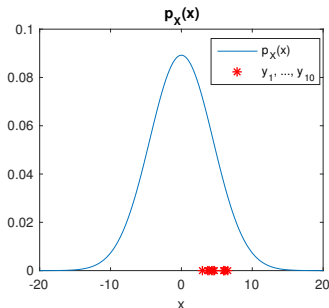
Bir başka Bayesci çıkarım örneği:

$$X \sim \mathcal{N}(\mu, \sigma^2), \quad Y_1, \dots, Y_n | X = x \stackrel{\text{i.i.d.}}{\sim} \text{Unif}(x - a, x + a).$$

Sonsal dağılım:

$$\pi(x) = p_{X|Y}(x|y) \propto \underbrace{\phi(x; \mu, \sigma^2)}_{p_X(x)} \underbrace{\frac{1}{(2a)^n} \prod_{i=1}^n \mathbb{I}_{(x-a, x+a)}(y_i)}_{p_{Y|X}(y|x)}$$

Öncül ve sonsal dağılımlar ( $n = 10$ ,  $a = 2$ ,  $\mu = 0$ ,  $\sigma^2 = 10$ ):



# Öz-düzgeleyici önem örnekleme: $q(x)$ 'in seçiminin önemi

Sonsal dağılım:

$$\pi(x) = p_{X|Y}(x|y) \propto \underbrace{\phi(x; \mu, \sigma^2)}_{p_X(x)} \underbrace{\frac{1}{(2a)^n} \prod_{i=1}^n \mathbb{I}_{(x-a, x+a)}(y_i)}_{p_{Y|X}(y|x)}$$

Öz-düzgeleyici önem örnekleme  $\mathbb{E}(X|Y=y)$ 'i kestirmek için kullanılabilir.

$q(x)$  için ilk seçim: öncül dağılım  $q(x) = \phi(x; \mu, \sigma^2)$ .

Bu geçerli bir seçimdir, ama  $a$  küçük ve  $\sigma^2$  büyük ise önem fonksiyonu

$$w(x) = \frac{1}{(2a)^n} \prod_{i=1}^n \mathbb{I}_{(x-a, x+a)}(y_i).$$

çoğu örnek için 0, çok az örnek için  $\frac{1}{(2a)^n}$  olacaktır. Bu da varyansın fazla olmasına sebep olur.



# Öz-düzgeleyici önem örnekleme: $q(x)$ 'in seçiminin önemi

Daha akıllı bir seçim sonsal dağılıma bakılarak yapılabilir.

$y_{\max} = \max_i y_i$  ve  $y_{\min} = \min_i y_i$  olsun.

$$x \in (y_{\max} - a, y_{\min} + a) \Leftrightarrow x - a < y_i < x + a, \quad \forall i = 1, \dots, n$$

Dolayısıyla,  $(y_{\max} - a, y_{\min} + a)$  aralığının dışında vakit harcamamıza gerek yok.

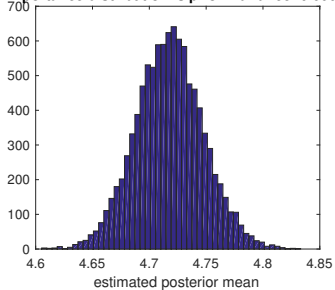
Mantıklı bir seçim:  $q(x) = \text{Unif}(x; y_{\max} - a, y_{\min} + a)$ .

Bu seçimle

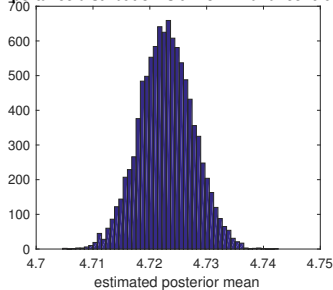
$$w(x) = \begin{cases} \frac{\phi(x; \mu, \sigma^2) \frac{1}{(2a)^n}}{1/(2a + y_{\min} - y_{\max})}, & x \in (y_{\max} - a, y_{\min} + a) \\ 0, & \text{else} \end{cases}$$

# Öz-düzgeleyici önem örnekleme: $q(x)$ 'in seçiminin önemi

importance distribution is prior: variance: 0.00089



importance distribution is uniform: variance: 0.00002



# 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  $\{X_t\}_{t \geq 1}$ :

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

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

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

# Metropolis-Hastings

# Markov zinciri Monte Carlo

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

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

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

Bu değerler,  $\pi$  dağılıma göre olan beklenti değerlerini hesaplamaya yarar.

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

# Metropolis-Hastings

En çok kullanılan MZMC yöntemlerinden biri *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.



# Metropolis-Hastings: doğruluk

Metropolis-Hastings'in geçiş matrisi (çekirdeği):

$$M(x'|x) = q(x'|x)\alpha(x, x') + \underbrace{\left[1 - \int_{\mathcal{X}} q(x'|x)\alpha(x, x')\right]}_{x' \text{ te reddetme olasılığı}} \delta_x(x').$$

$\pi$ ,  $M$  için ayrıntılı denge koşulunu sağlar: Herhangi  $A, B \subseteq \mathcal{X}$  kümeleri için

$$\int_A \int_B M(x'|x)\pi(x)dx dx' = \int_B \int_A M(x'|x)\pi(x)dx dx'$$

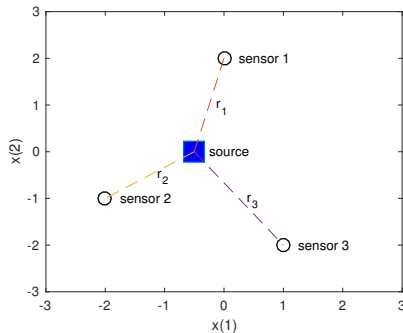
( $X$  ayrıkça, her  $x, x'$  için  $M(x'|x)\pi(x) = M(x|x')\pi(x')$ .)

$\pi$ ,  $M$  için ayrıntılı denge koşulunu sağlarsa,

- ▶  $M$  tersinebilirdir ve
- ▶  $\pi$ ,  $M$ 'in değişimsiz dağılımıdır.

# 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:

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

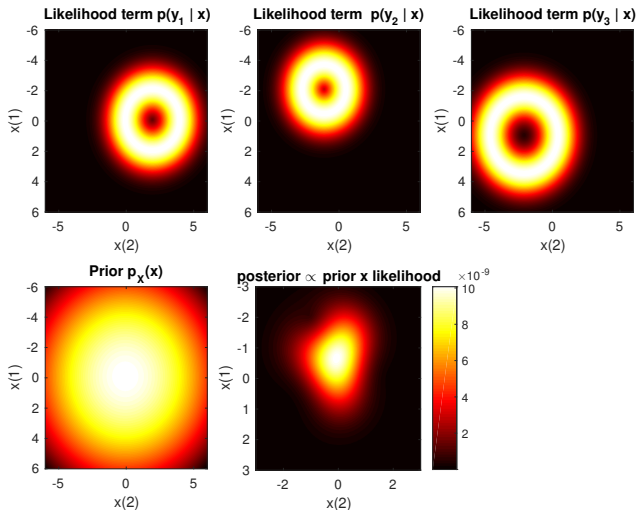
Amaç:  $Y = y = (y_1, y_2, y_3)$  verildiğinde  $p_{X|Y}(x|y)$ 'i bulmak ve  $\mathbb{E}(X|Y = y)$ 'yi hesaplamak.

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

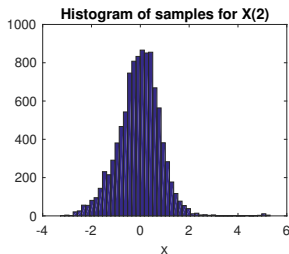
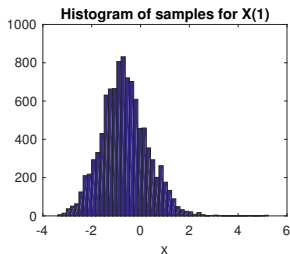
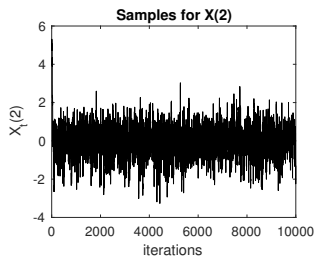
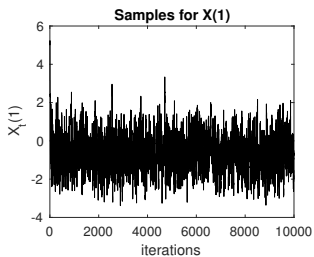
Metropolis-Hastings algoritması  $q(x'|x) = \phi(x'; x, \sigma_q^2 I_2)$  alınarak yapılabilir:

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

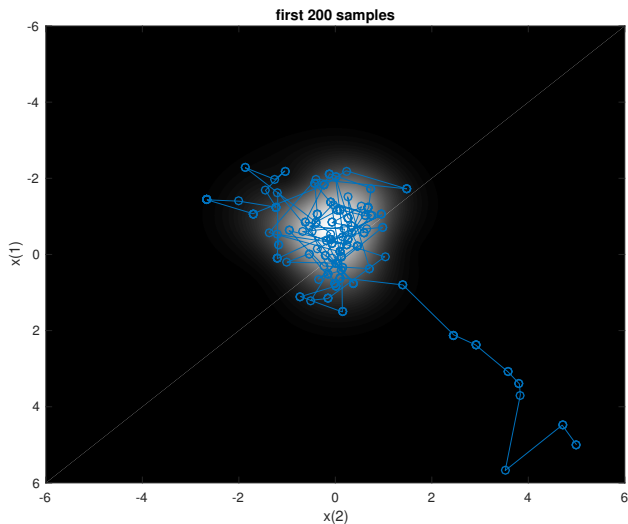
# Hedef yer saptaması - dağılımlar, $\sigma_x^2 = 100$ , $\sigma_y^2 = 1$



# Hedef yer saptaması için Metropolis-Hastings: örnekler



# Hedef yer saptaması için Metropolis-Hastings: ilk 200 örnek



# Gibbs örneklemesi

# Gibbs örneklemesi

Bir diğer çok sık kullanılan MZMC yöntemi de *Gibbs örneklemesidir*.

Uygulanması için

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

**Gibbs örneklemesi:**

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

$n = 2, 3, \dots$  için,

$k = 1, \dots, d$  için

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



# Gibbs örneklemesi: doğruluk

## Gibbs örneklemesi:

$X_1 = (X_1(1), \dots, X_1(d))$  ile başla.  $n = 2, 3, \dots$  için,  
 $k = 1, \dots, d$  için

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

Bu algoritmanın döngüsünün  $k$ 'inci adımına karşılık gelen geçiş matrisi (çekirdeği)  $M_k$ :

$$M_k(x, y) = \pi_k(y_k | x_{-k}) \delta_{x_{-k}}(y_{-k})$$

Burada  $x_{-k} = (x_{1:k-1}, x_{k+1:d})$   $y_{-k} = (y_{1:k-1}, y_{k+1:d})$ .

Bütün bir döngüye karşılık gelen geçiş matrisi (çekirdeği)

$$M = M_1 M_2 \dots M_d$$

Her bir  $M_k$ 'nin ayrıntılı denge koşulunu sağladığı ve  $\pi$ 'ye göre tersinirliği gösterilebilir.

# 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  $\varphi_n : \mathcal{X}^n \rightarrow \mathbb{R}$  olan bir  $\{\varphi_n\}_{n \geq 1}$  fonksiyon dizisi verilsin.

Amaç: Sıralı çıkarım

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

integrallerine sıralı bir şekilde nasıl yaklaşabiliriz?

# Örnek: Saklı Markov modelleri (SMM)

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

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

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

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

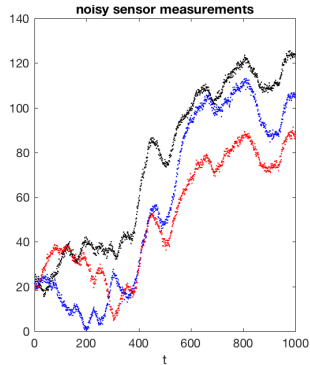
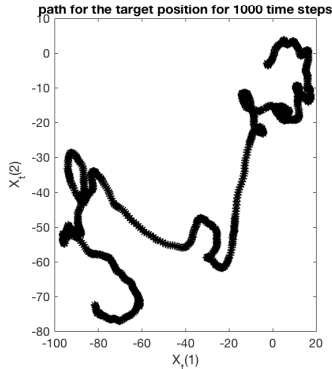
$\{Y_t\}_{t \geq 1}$ ,  $\{X_t\}_{t \geq 1}$ 'ye koşullu bağımsız süreç:

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

# Örneğe örnek: Hedef takip

- ▶  $V_t = (V_t(1), V_t(2))$ :  $t$  anındaki hız vektörü
- ▶  $P_t = (P_t(1), P_t(2))$ :  $t$  anındaki pozisyon vektörü
- ▶  $X_t = (V_t, P_t)$ :  $t$  anındaki hız ve pozisyon
- ▶  $Y_t \sim \mathcal{N}(\|S_1 - P_t\|, \|S_2 - P_t\|, \|S_3 - P_t\|, \sigma_y^2 I_3)$ : 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 dağılım:

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

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 örnekleme

$\mathbb{E}_{\pi_n} [\varphi_n(X_{1:n})]$  için önem örnekleme 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:

$$q_n(x_{1:n}) = q_1(dx_1) \prod_{i=1}^n q_i(x_i | x_{1:i-1})$$

Bu durumda ağırlık fonksiyonları özyinelemeli olarak yazılabilir:

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

$\pi_n(x_{1:n}) = \hat{\pi}_n(x_{1:n}) / Z_n$  ve  $\hat{\pi}_n(x_{1:n})$  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 örnekleme

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

Öz-düzgeleyici önem örnekleme sıralı bir şekilde uygulanabilir:

$n = 1, 2, \dots$  için;

►  $i = 1, \dots, N$  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

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

► Öz-düzgeleyici önem ağırlıkları:  $i = 1, \dots, N$  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:  $\pi_n(x_{1:n}) \propto \hat{\pi}_n(x_{1:n}) = p(x_{1:n}, y_{1:n})$

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

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

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

## Parçacık süzgeci

# Ağırlık bozulması sorunu

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

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

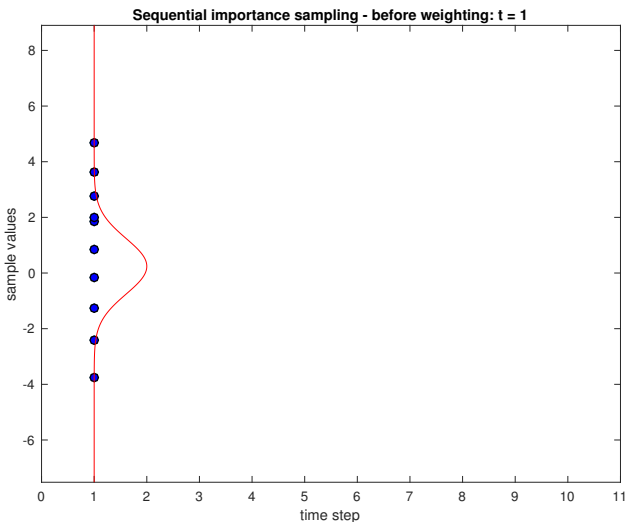
Limitte,  $W_n^{(i)}$ 'lerden bir tanesi 1, diğerleri 0 olacaktır.

Bu soruna, ağırlık bozulması sorunu denir.

# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

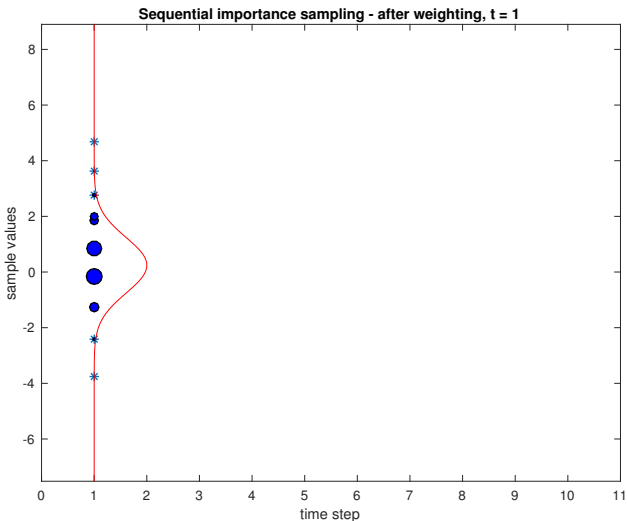
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

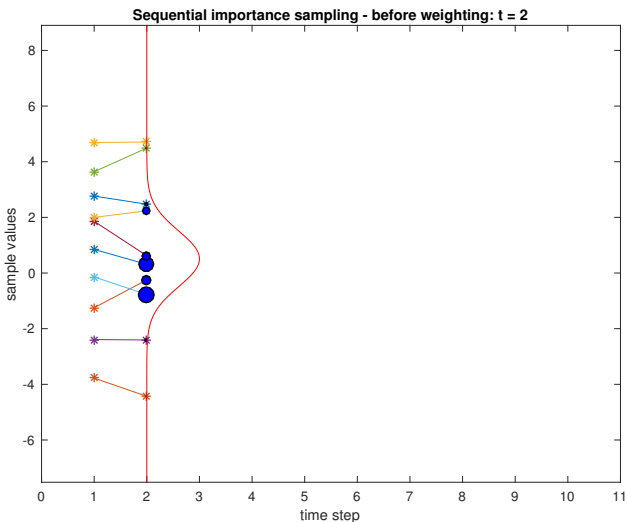
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

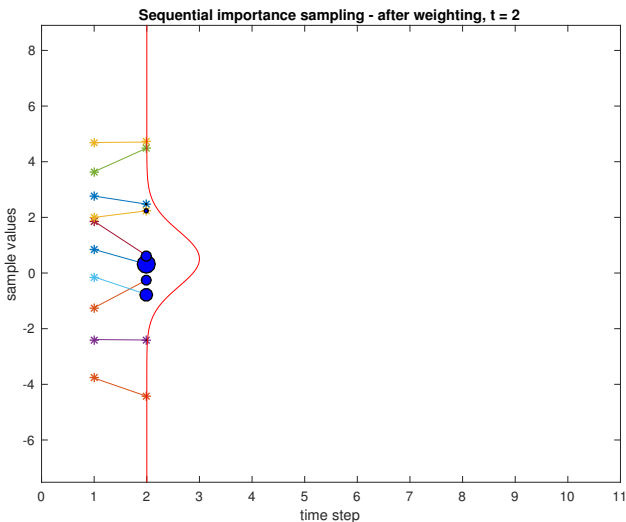
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

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

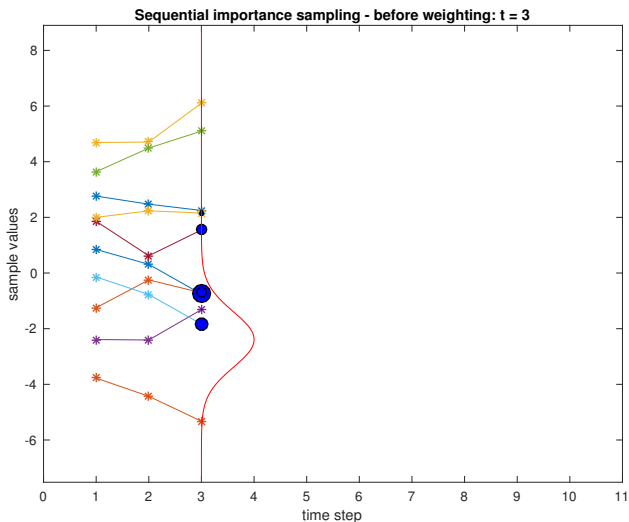




# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

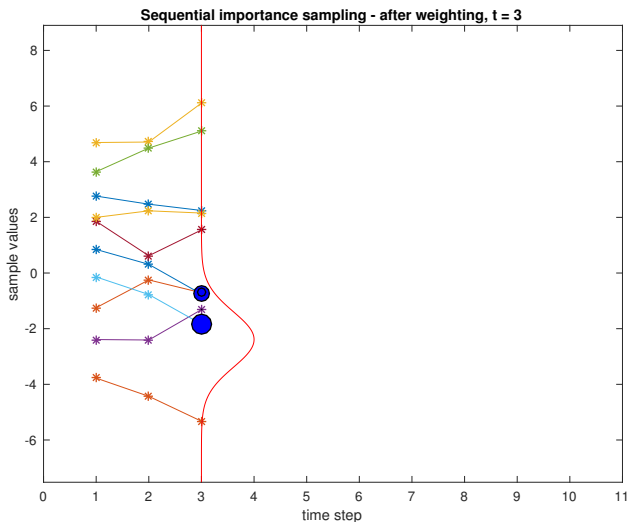
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

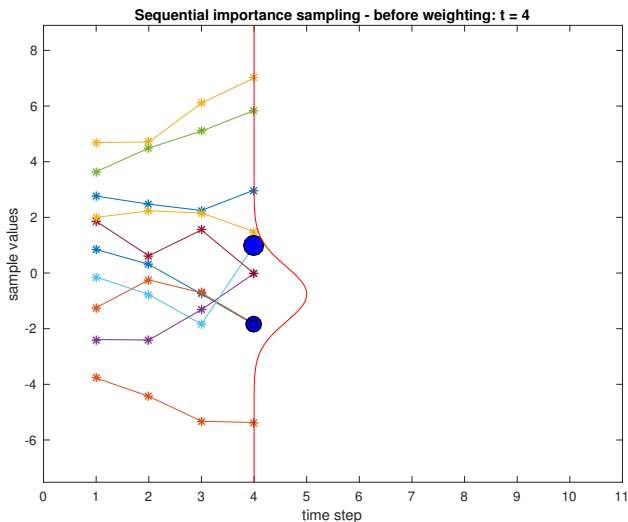
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

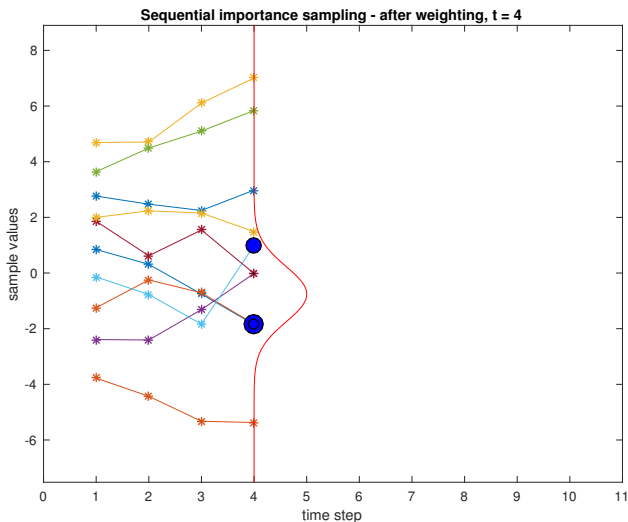
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

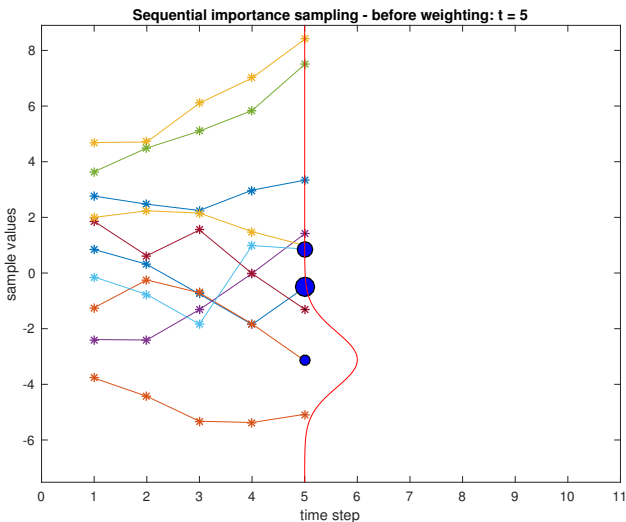
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

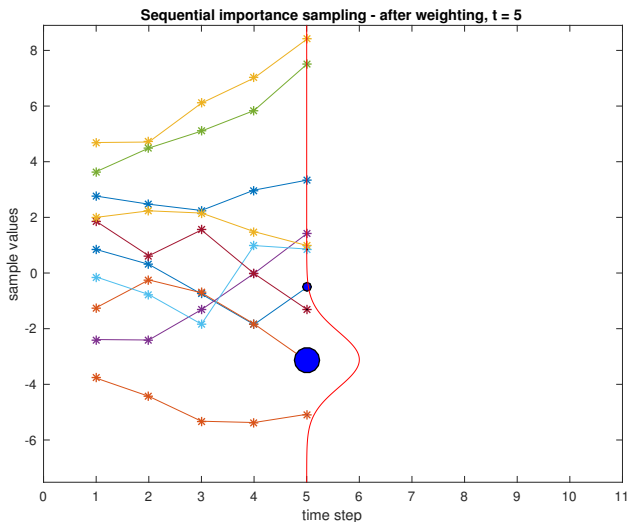
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

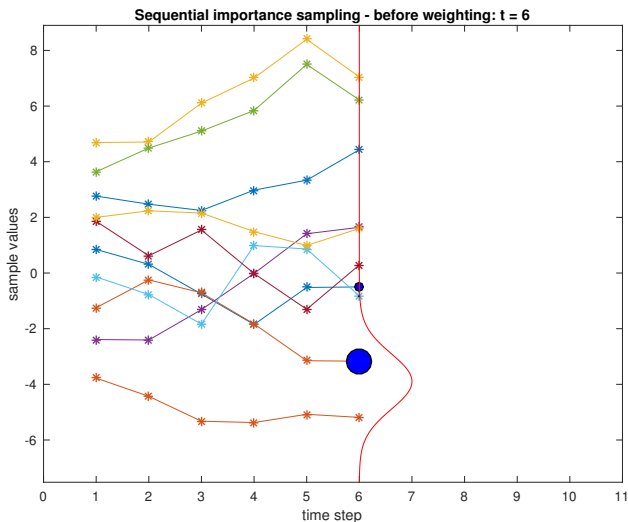
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

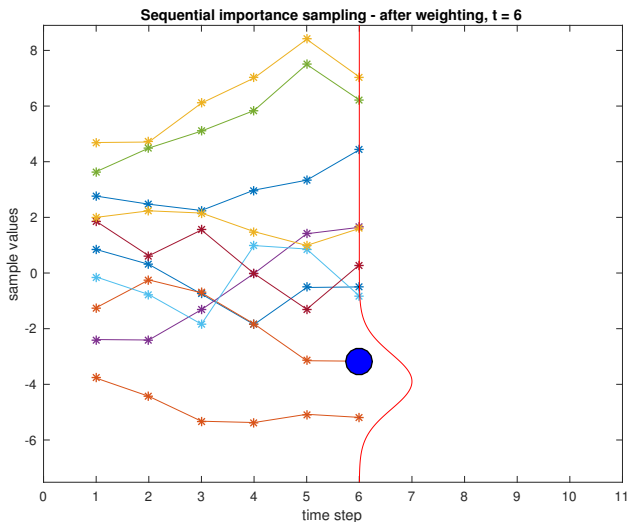
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

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

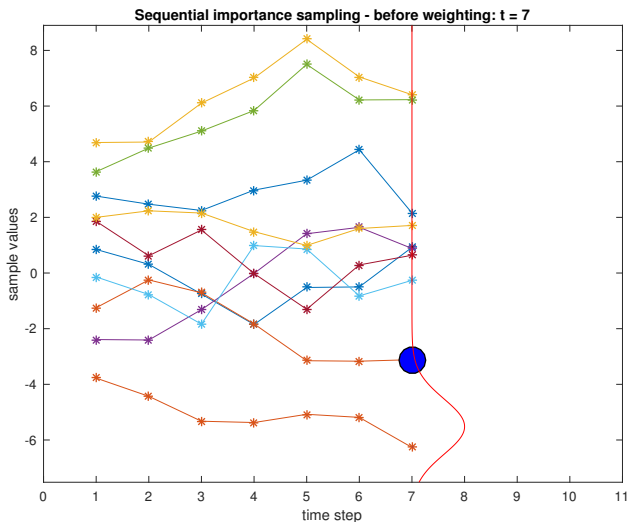




# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

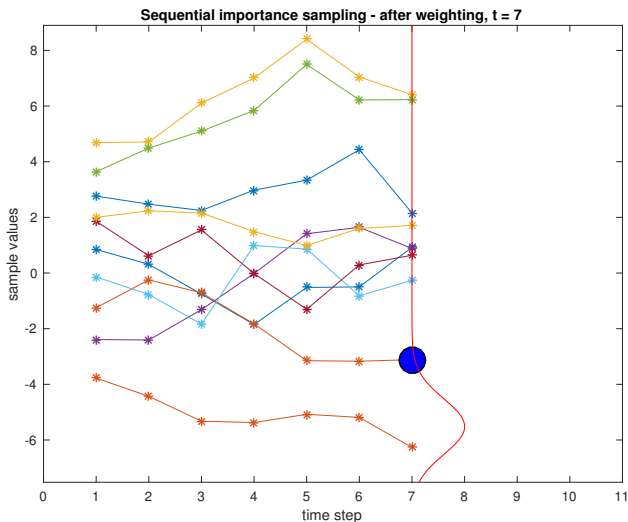
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

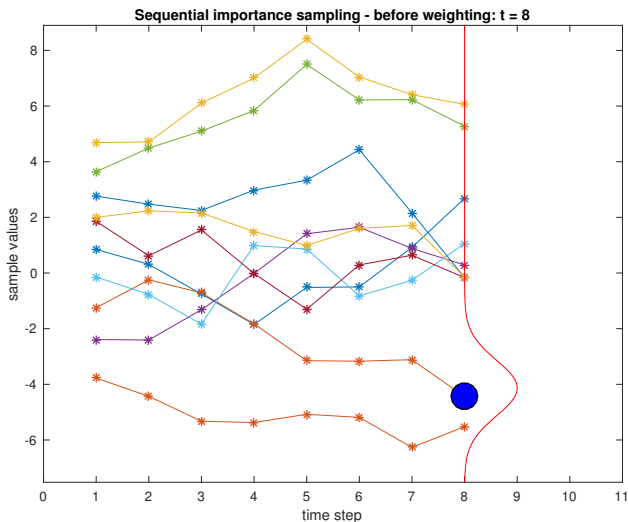
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

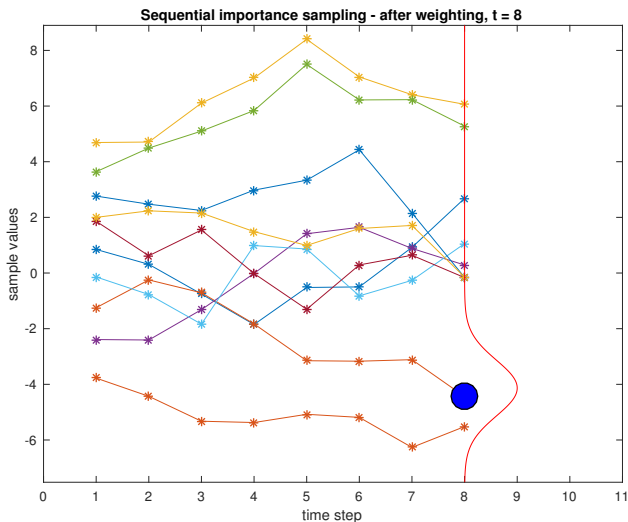
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

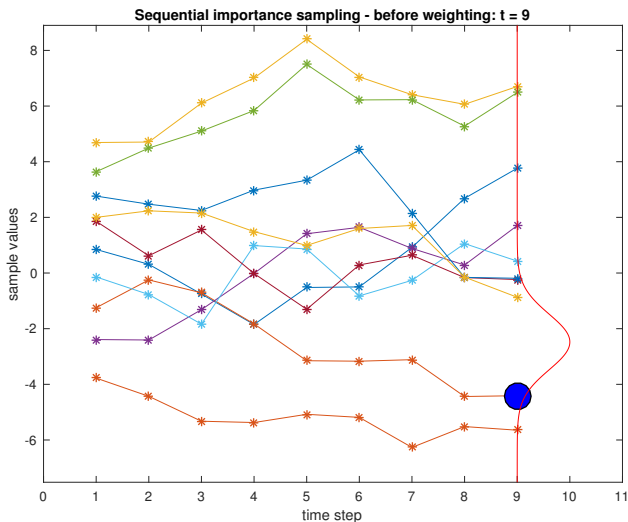
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

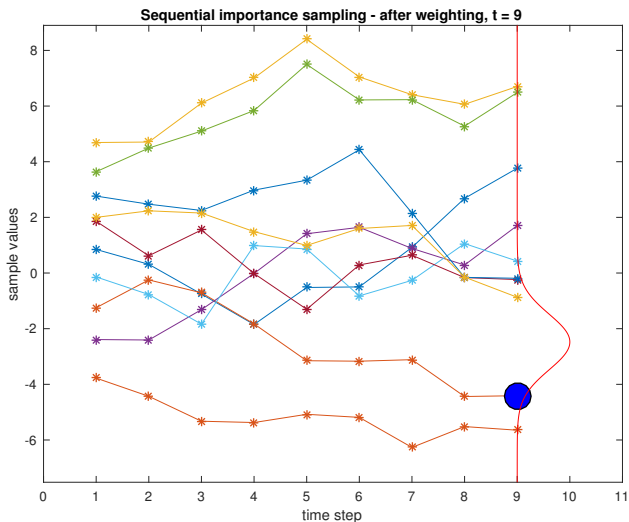
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

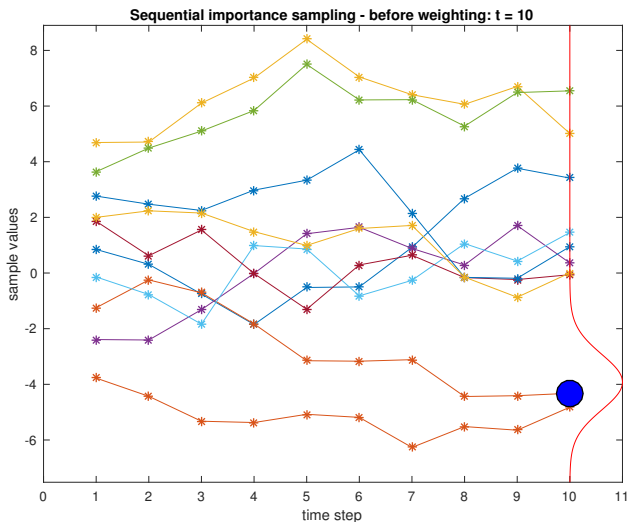
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# Ağırlık bozulması sorunu: Örnek

Doğrusal Gauss saklı Markov modeli

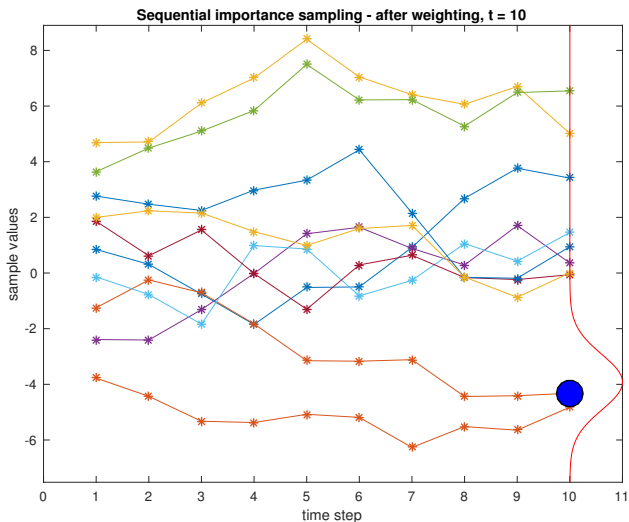
$$\eta(x) = \phi(x; 0, \sigma_0^2), \quad f(x'|x) = \phi(x'; ax, \sigma_x^2), \quad g(y|x) = \phi(y; bx, \sigma_y^2)$$



# 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 $\rightarrow$ Parçacık süzgeci

Ağırlık bozulması sorununu önlemek için, sıralı önem örneklemesinin her bir adımına yeniden örnekleme uygulanır.

**Yeniden örnekleme:** Ağırlıkları olan bir örnek kümesinin, yine o kümeden ağırlıklarına doğru orantılı ihtimallerle seçilmiş eşit ağırlıklandırılan örnekler kümesiyle değiştirilmesi.

Diyelim ki  $n - 1$  zamanında  $X_{1:n-1}^{(1)}, \dots, X_{1:n-1}^{(N)}$  örneklerimiz var ve bunların öz-düzgelenmiş ağırlıkları  $W_{n-1}^{(1)}, \dots, W_{n-1}^{(N)}$ .

Bu örneklerden, ağırlıkları olasılıklarıyla  $N$  kere bağımsız örnekler çekilir:

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

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

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

# SMM için parçacık süzgeci

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

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

**Parçacık süzgeci:**

$n = 1$  için;

$i = 1, \dots, N$  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.

$n = 2, 3, \dots$  için,

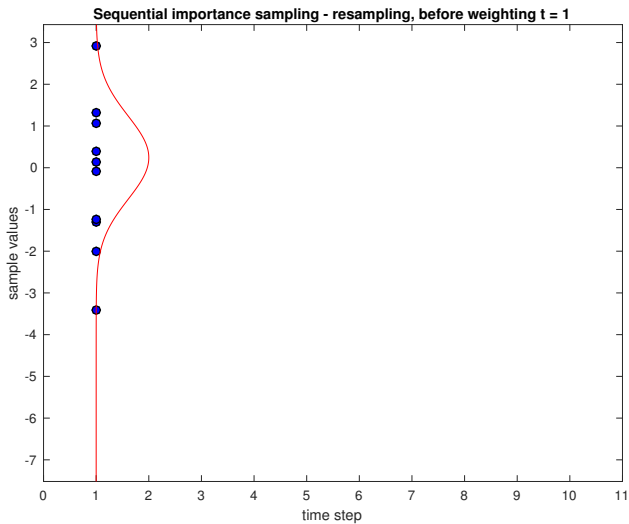
- ▶ Yeniden örnekleme ile  $\tilde{X}_{1:n-1}^{(1)}, \dots, \tilde{X}_{1:n-1}^{(N)}$  üretilir:

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

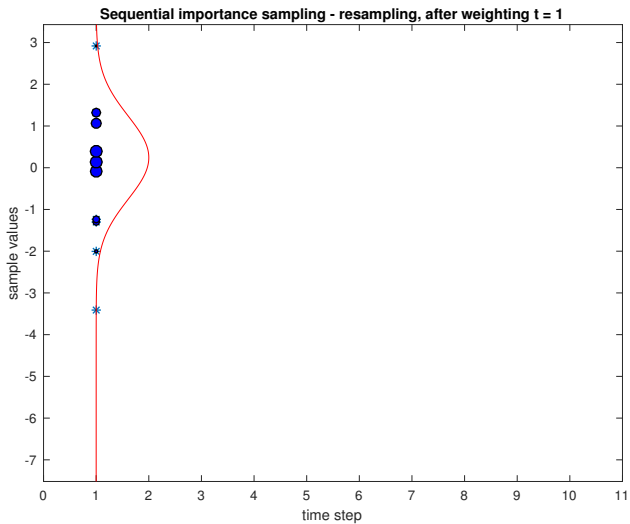
- ▶  $i = 1, \dots, N$  için,  $X_n^{(i)} \sim q_n(\cdot|\tilde{X}_{n-1}^{(i)}, y_n)$  örneklenir,  $X_{1:n}^{(i)} = (\tilde{X}_{1:n-1}^{(i)}, X_n^{(i)})$  oluşturulur.
- ▶ Bu parçacıkların ağırlıkları

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

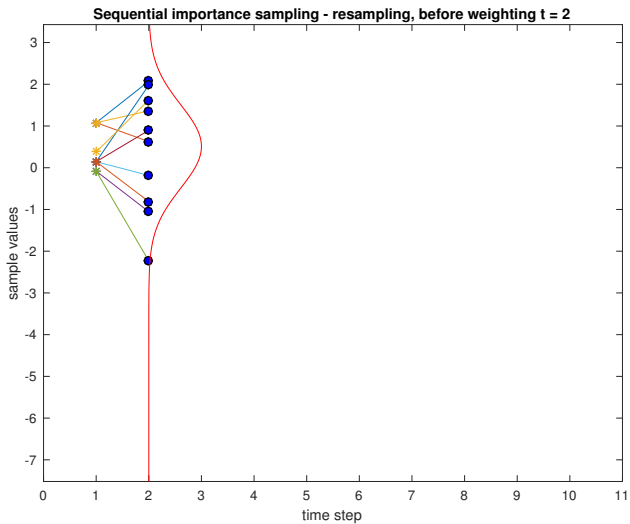
# Parçacık süzgeci



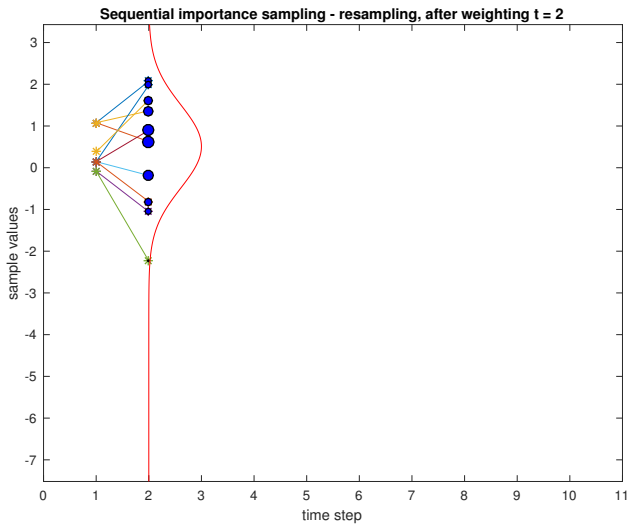
# Parçacık süzgeci



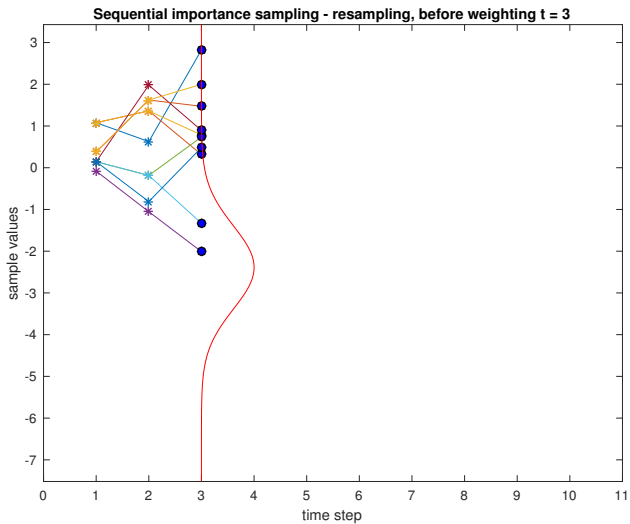
# Parçacık süzgeci



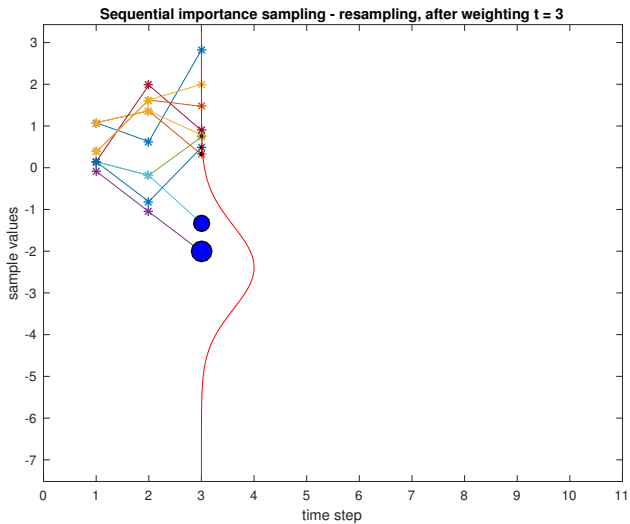
# Parçacık süzgeci



# Parçacık süzgeci

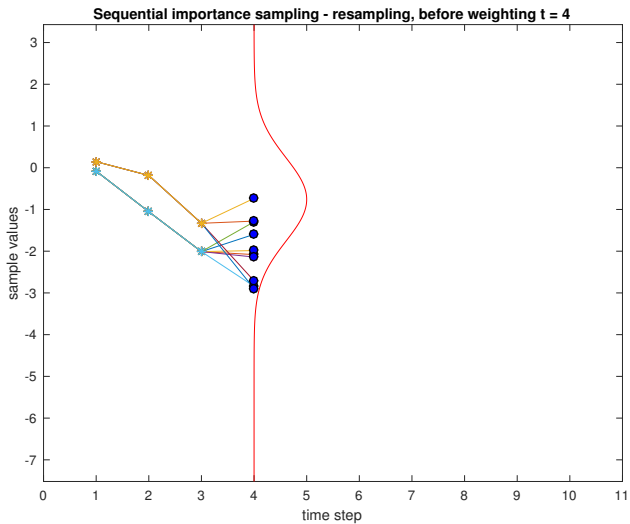


# Parçacık süzgeci

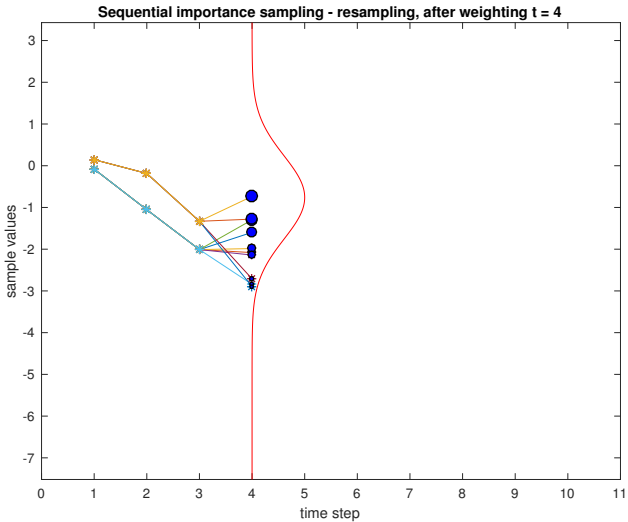




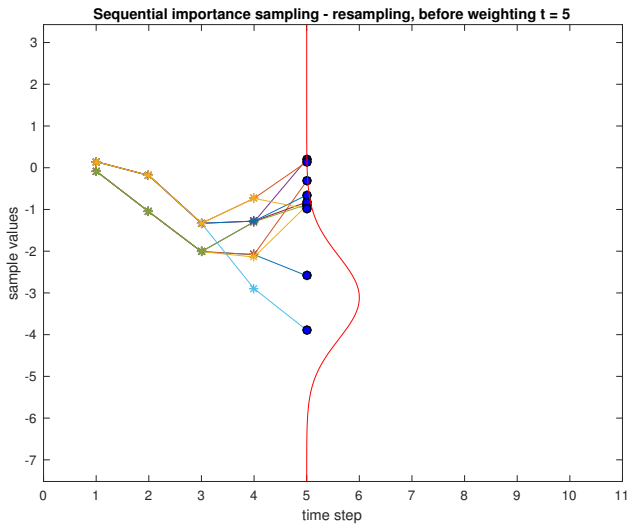
# Parçacık süzgeci



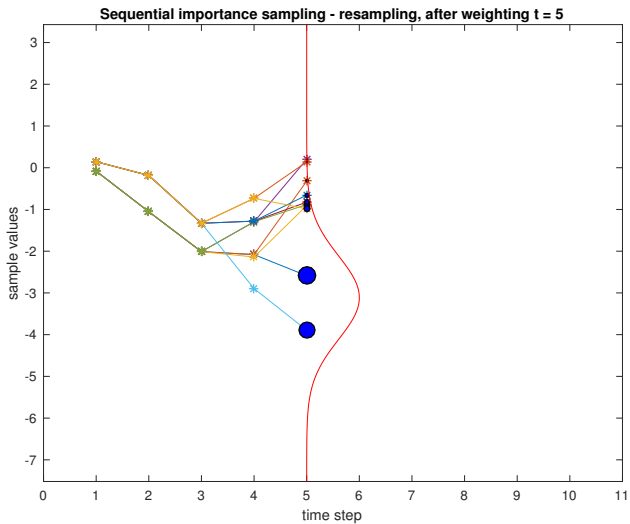
## Parçacık süzgeci



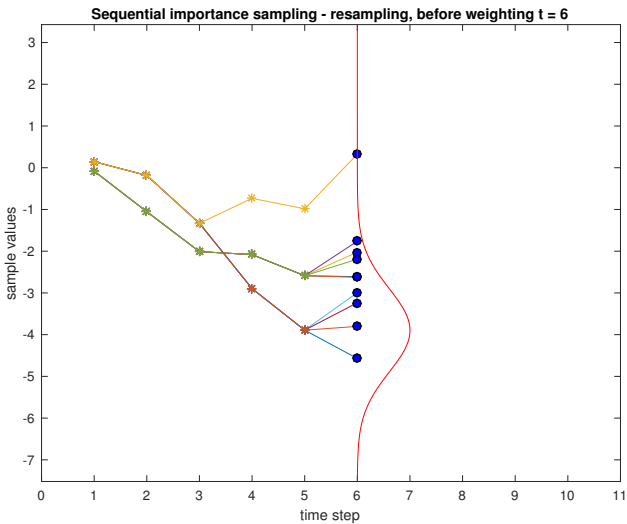
# Parçacık süzgeci



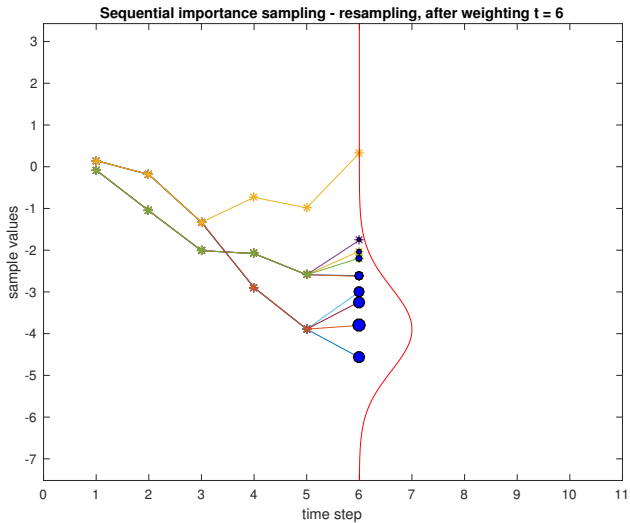
# Parçacık süzgeci



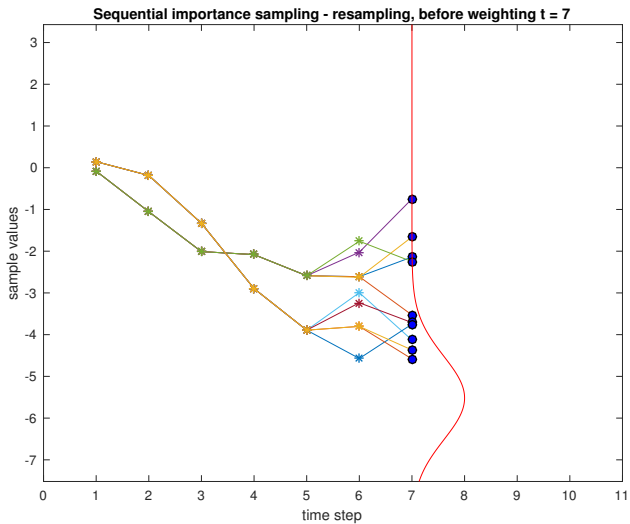
# Parçacık süzgeci



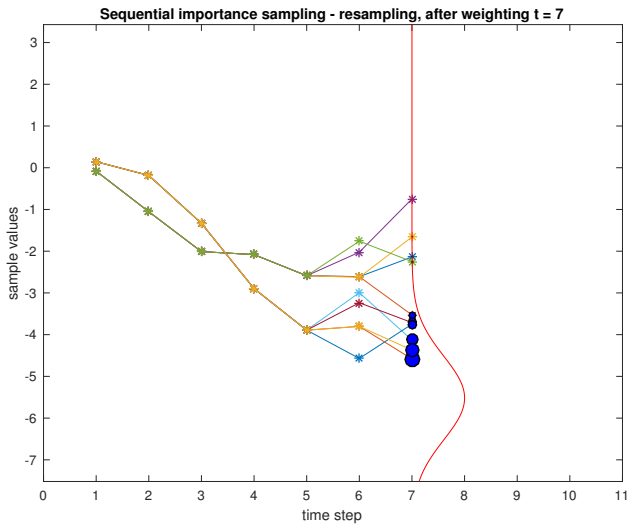
# Parçacık süzgeci



# Parçacık süzgeci

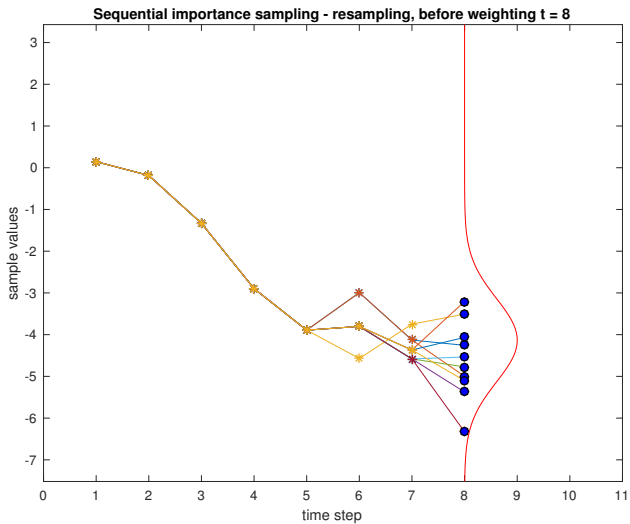


# Parçacık süzgeci

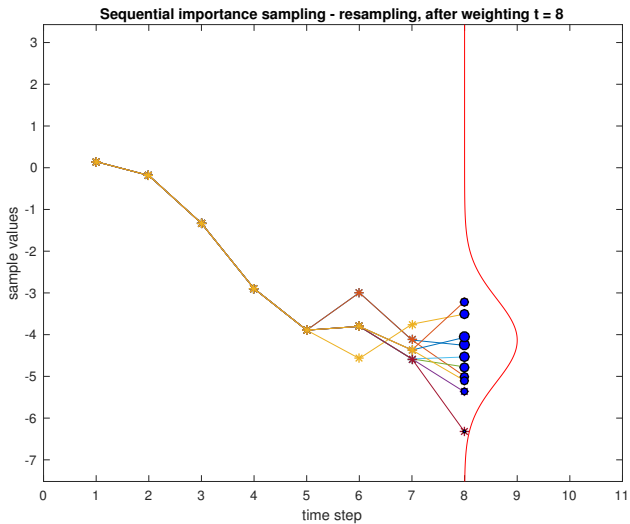




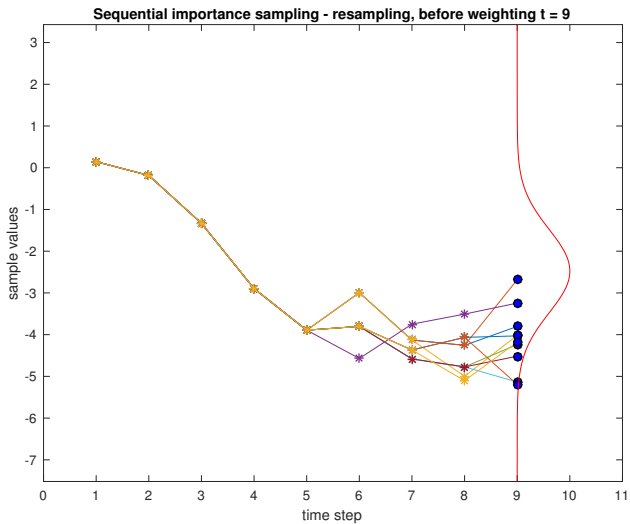
# Parçacık süzgeci



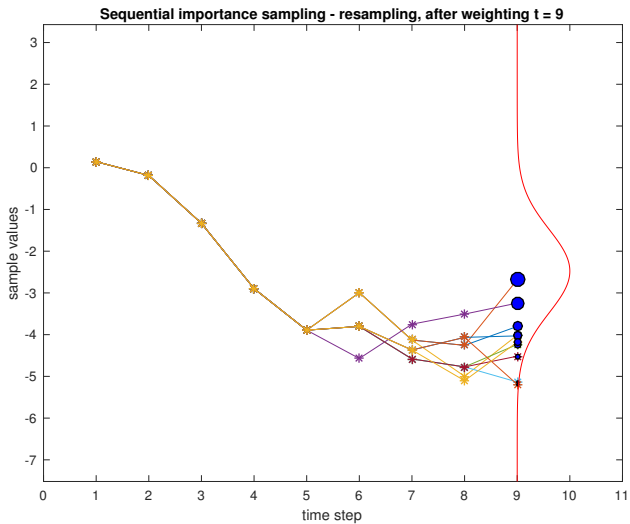
# Parçacık süzgeci



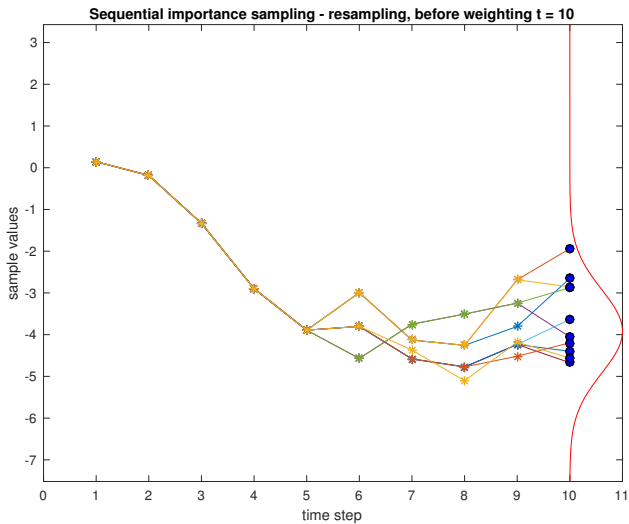
# Parçacık süzgeci



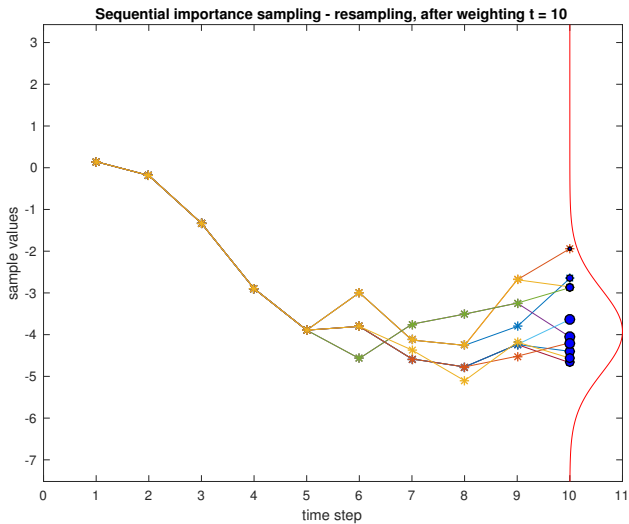
# Parçacık süzgeci



# 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 örnekleme sebebiyle önceki zamanlara ait parçacık sayısı gitgide düşer.