

Monte Carlo Yöntemleri

Sinan Yıldırım

MDBF, Sabancı Üniversitesi

7 Eylül 2019

Giriş

Bayes Teoremi



$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

İki zar atılıyor, gelen sayılar X_1 ve X_2 olsun.

$S = X_1 + X_2 = 8$ ise en az bir zardan 3 gelmiş olma olasılığı nedir?

$$A = \{(X_1, X_2) : X_1 = 3 \text{ or } X_2 = 3\}.$$

$$B = \{(X_1, X_2) : S = 8\}.$$

$$B = \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}, \quad A \cap B = \{(3, 5), (5, 3)\}.$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{2/36}{5/36} = \frac{2}{5}.$$

İki zar atılıyor, gelen sayılar X_1 ve X_2 olsun.

Toplam çift olduğu bilindiğine göre, toplamın en fazla 4 olma olasılığı nedir?

$$A = \{(X_1, X_2) : X_1 + X_2 \leq 4\},$$

$$B = \{(X_1, X_2) : X_1 + X_2 \text{ çift}\}$$

$$B = \{(X_1, X_2) : X_1, X_2 \text{ çift}\} \cup \{(X_1, X_2) : X_1, X_2 \text{ tek}\}.$$

$$A \cap B = \{(1, 1), (1, 3), (3, 1), (2, 2)\}.$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{4/36}{3/6 \times 3/6 + 3/6 \times 3/6} = \frac{2}{9}.$$

Model seçimi

Bir torbada iki para var: Biri hilesiz, diğerinin yazı gelme olasılığı $\rho = 0.3$.

Biri torbadan rastgele bir para seçiyor ve bunu 10 defa atıyor.

$$\mathcal{D} = (T, Y, T, T, Y, T, T, T, Y, T)$$

Hilesiz paranın çekilmiş olma olasılığı nedir?

İki modelimiz var:

- ▶ M_1 : hilesiz para çekildi,
- ▶ M_2 : hileli para çekildi.

$$\begin{aligned}\mathbb{P}(M_1|\mathcal{D}) &= \frac{\mathbb{P}(\mathcal{D}, M_1)}{\mathbb{P}(\mathcal{D})} = \frac{\mathbb{P}(M_1)\mathbb{P}(\mathcal{D}|M_1)}{\mathbb{P}(\mathcal{D}|M_1)\mathbb{P}(M_1) + \mathbb{P}(\mathcal{D}|M_2)\mathbb{P}(M_2)} \\ &= \frac{1/2^{10}}{1/2^{10} + \rho^{n_T}(1 - \rho)^{n_H}} = 0.3052.\end{aligned}$$

DNA testi

Bir suç mahalinde DNA izine rastlanıyor. Şüphelilerden DNA örneği alınıp bulunan DNA ile karşılaştırılıyor.

Testin performansı:

	Pozitif	Negatif
Suçlu	1	0
Suçsuz	0.001	0.999

Bir X kişisi teste tabi tutuluyor ve testin sonucu pozitif çıkıyor.

X kişisi suçlu mu?

Rassal değişkenler ve Bayes teoremi

Ayrık değişkenler için:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} = \frac{p(x)p(y|x)}{\sum_x p(x,y)}$$

Sürekli değişkenler için:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)} = \frac{p(x)p(y|x)}{\int_x p(x,y)dx}$$

$p(x, y)$	$y = 1$	$y = 2$	$y = 3$	$y = 4$	$p(x)$
$x = 1$	$1/20$	$3/20$	$4/20$	$2/20$	
$x = 2$	$1/20$	$3/20$	$3/20$	$3/20$	
$p(y)$					

$p(x y)$	$y = 1$	$y = 2$	$y = 3$	$y = 4$
$x = 1$				
$x = 2$				

Konjuge öncül dağılımlar

sonsal dağılım \propto öncül dağılım \times olabilirlik.

Çoğu zaman, olabilirlik $p(y|x)$ veriyi üreten süreç tarafından belirlenir.

Seçme özgürlüğü daha çok öncül dağılım $p(x)$ için söz konusudur.

Bazı durumlarda $p(x)$ öyle bir seçilebilir ki, $p(x|y)$, $p(x)$ ile aynı olasılık dağılım ailesinin başka bir üyesi olur.

Bu tür öncül dağılımlara olabilirlik için konjuge dağılımlar denir.

Binom dağılımının parametresi

Hileli bir paranın tura ihtimali $X = \lambda = \mathbb{P}(T)$.

Öncül dağılım: $X = \lambda \sim \text{Beta}(a, b)$

Para n kere atılıyor. Turların sayısı

$$Y \sim \text{Binom}(n, \lambda)$$

Para $Y = k$ kere tura geliyor. X 'in sonsal dağılımı nedir?

Sonsal dağılım:

$$p(x|y) \propto p(x)p(y|x) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}$$

$p(x|y) \propto p(x, y)$ 'dan hareketle

$$p(x|y) \propto x^{a+k-1}(1-x)^{b+n-k-1}$$

Bu dağılım başka bir Beta dağılımıdır:

$$X|Y = k \sim \text{Beta}(a_{x|y}, b_{x|y})$$

öyle ki, $a_{x|y} = a + k$ and $b_{x|y} = b + n - k$.

Normal dağılımın lokasyon parametresi

Olabilirlik:

$$Y_1, \dots, Y_n \sim \mathcal{N}(x, \sigma^2)$$

Öncül dağılım:

$$X \sim \mathcal{N}(0, \sigma_x^2).$$

Sonsal dağılım:

$$X | Y_{1:n} = y_{1:n} \sim \mathcal{N}(\mu_{x|y}, \sigma_{x|y}^2)$$

$$\mu_{x|y} = \left(\frac{1}{\sigma_x^2} + \frac{n}{\sigma^2} \right)^{-1} \frac{1}{\sigma^2} \sum_{i=1}^n y_i, \quad \sigma_{x|y}^2 = \left(\frac{1}{\sigma_x^2} + \frac{n}{\sigma^2} \right)^{-1}$$

Normal dağılımın varyans parametresi

Olabilirlik: normal dağılım

$$Y_1, \dots, Y_n \sim \mathcal{N}(\mu, x)$$

Öncül dağılım: Ters gamma

$$X \sim \mathcal{IG}(\alpha, \beta)$$

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right).$$

$$X | Y_{1:n} = y_{1:n} \sim \mathcal{IG}(\alpha_{x|y}, \beta_{x|y})$$

$$\alpha_{x|y} = \alpha + \frac{n}{2}, \quad \beta_{x|y} = \beta + \frac{1}{2} \sum_{i=1}^n (y_i - \mu)^2.$$

$p(y)$: Kanıt

Konjuge öncüller kullanıldığında, $p(y)$ de hesaplanabilir.

$$p(y) = \frac{p(x, y)}{p(x|y)} = \frac{p(x)p(y|x)}{p(x|y)}$$

Örnek: Tura olasılığı parametresi $X \sim \text{Beta}(\alpha, \beta)$ olan bir para n kere atılıyor. k tura gelme olasılığı nedir?

$$\begin{aligned} p(k) &= \frac{\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}}{\frac{x^{\alpha+k-1}(1-x)^{\beta+n-k-1}}{B(\alpha+k, \beta+n-k)}} \\ &= \frac{n!}{k!(n-k)!} \frac{B(\alpha+k, \beta+n-k)}{B(\alpha, \beta)}. \end{aligned}$$

Bu dağılım Beta-Binomial dağılımı olarak da bilinir.

Sonsal beklenti değeri

$Y = y$ verildiğinde,

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

Herhangi bir rassal $\hat{X}(Y)$ kestirimi için, **ortalama kare hata**:

$$\mathbb{E}([X - \hat{X}(Y)]^2) = \int (x - \hat{X}(y))^2 p(x, y) dx dy$$

$\hat{X}(Y) = \mathbb{E}(X|Y)$, ortalama kare hatayı en küçük yapar.

Sonsal olasılığı en büyük yapan değer

$$\hat{x}_{\text{MAP}} = \arg \max_{x \in \mathcal{X}} p(x|y) = \arg \max_{x \in \mathcal{X}} p(x, y).$$

Enbüyük olabilirlik kestirimine benzer:

$$\hat{x}_{\text{MLE}} = \arg \max_{x \in \mathcal{X}} p(y|x),$$

fakat ilkinde $p(x)$ de hesaba katılır.

Sonsal dağılımdan gözlem tahminine

Model: y_1, y_2, \dots değişkenleri, x değişkenine koşullu bağımsız.

Şimdiye kadarki veri: y_1, \dots, y_n .

Amaç: y_{n+1} 'in tahmini.

Olası bir yöntem (klasik istatistik):

- ▶ x 'in noktasal kestirimini elde et: \hat{x} .

Örn: $\hat{x}_{\text{MLE}}, \hat{x}_{\text{MAP}}$

- ▶ $p(y_{n+1}|\hat{x})$ dağılımından yola çıkarak y_{n+1} 'i tahmin et.

Bayesçi yöntem: $p(y_{n+1}|y_{1:n})$ 'yi hesapla:

$$\begin{aligned} p(y_{n+1}|y_{1:n}) &= \int p(y_{n+1}, x|y_{1:n}) dx \\ &= \int p(y_{n+1}|x) \underbrace{p(x|y_{1:n})}_{\text{sonsal dağılım}} dx \end{aligned}$$

Bernoulli deneyleri

Her biri $\mathcal{B}(x)$ olan n denemeden $Y = k$ tanesi başarılı.

Öncül dağılım: $X \sim \text{Beta}(\alpha, \beta)$.

Bir sonraki m denemeden $Z = r$ tanesinin başarılı olma olasılığı nedir?

X in $Y = k$ verildiğinde sonsal dağılımı:

$$\text{Beta}(\alpha' = \alpha + k, \beta' = \beta + n - k)$$

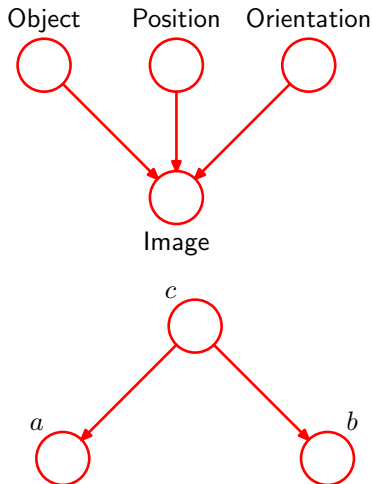
$$\begin{aligned}\mathbb{P}(Z = r | Y = k) &= \int_0^1 \text{Binom}(m; r, x) \text{Beta}(x; \alpha', \beta') dx \\ &= \frac{m!}{r!(m-r)!} \frac{B(\alpha' + r, \beta' + m - r)}{B(\alpha', \beta')}.\end{aligned}$$

Çizge modelleri

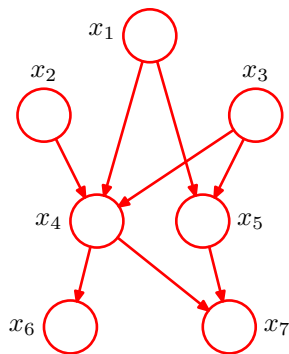
Bu bölümdeki örnek çizgeler, Christopher M. Bishop'un "Pattern Recognition and Machine Learning" adlı kitabından alınmıştır.

Yönlü düz çizgeler (Bayes ağları)

Rassal değişkenlerin arasındaki koşullu bağımsızlık ilişkilerini sebep sonuç ilişkisini temel alarak verir.

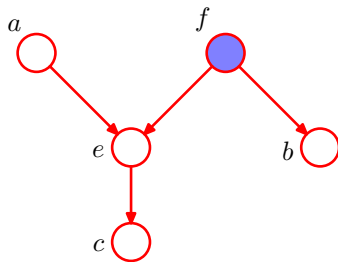
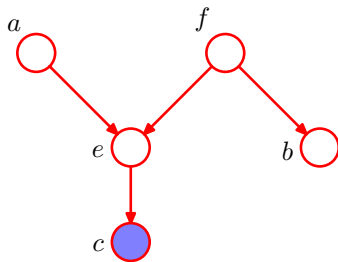


Yönlü düz çizgeler - Olasılık dağılımı



$$p(x_{1:7}) = p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2, x_3)p(x_5|x_1, x_3)p(x_6|x_4)p(x_7|x_4, x_5)$$

Koşullu bağımsızlık



► $a \perp b | c$ midir?

► $a \perp b | f$ midir?

Koşullu bağımsızlık - d -ayrılık

A , B ve C düğüm kümeleri olsun.

A ve B C 'ye koşullu bağımsız mıdır?

$$A \perp B | C$$

d -ayrılık

A 'daki bir düğümden B 'deki bir düğüme giden bir yol ele alalım.

Bu yolun üstündeki herhangi bir düğümü ele alalım:

- Bu düğüme ulaşan oklar **kuyruk-kuyruğa** ve **kafa-kuyruğa** ise ve bu düğüm C 'nin içindeyse; veya
- Bu düğüme gelen oklar **kafa-kafaya** ise ve ne bu düğüm ne de onun alt-düğümü C 'nin içinde ise (neither-nor);

bu yol C tarafından engellenmiş sayılır.

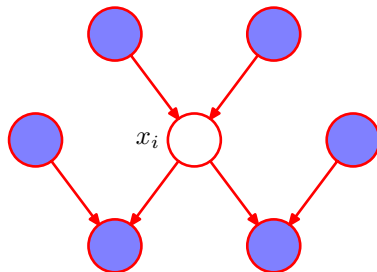
Eğer A 'daki her bir düğümden B 'deki her bir düğüme giden bütün yollar C tarafından engellenmişse, A ve B , C tarafından d -ayrılmıştır, ve

$$A \perp B | C$$

sağlanır.

Markov battaniyesi

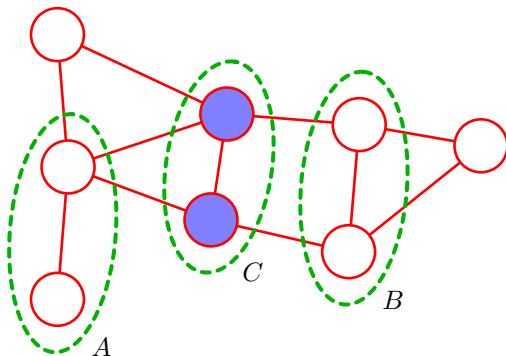
- ▶ Bir üst-düğüm (ebeveynler)
- ▶ bir alt-düğüm (çocuklar),
- ▶ çocukların beraber yapıldığı partnerler



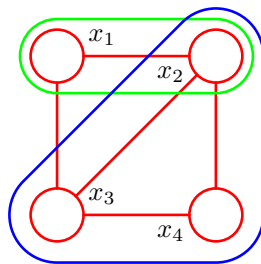
Bir düğüm, Markov battaniyesi verildiğinde diğer düğümlerden bağımsızdır.

Yönsüz çizgeler (Markov rastgele alanı)

Koşullu bağımsızlık özelliklerini, değişkenler arasındaki ilintilere dayanarak verir.

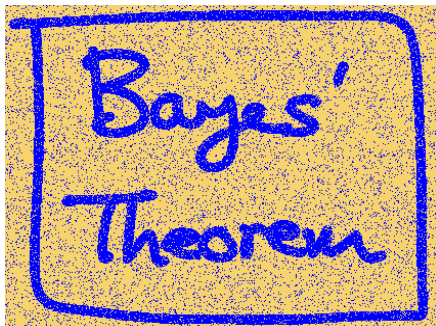
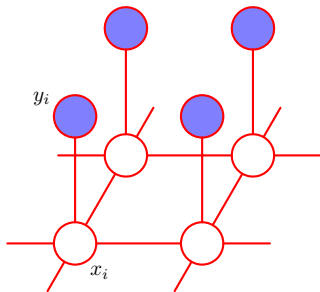


Klikler, potansiyeller ve birleşik dağılım



$$p(x_{1:4}) = \frac{1}{Z} \psi_1(x_1, x_2) \psi_2(x_2, x_3, x_4).$$

İmge işleme örneği



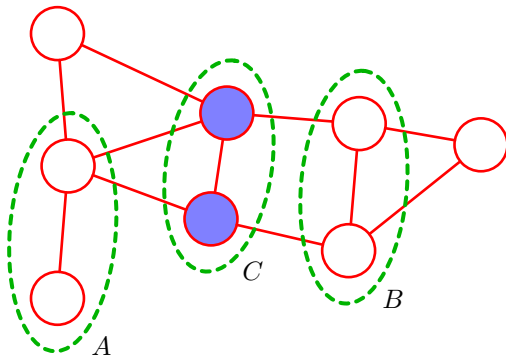
y : gürültülü imge, x : asıl imge (bilinmiyor)

$$x_i \in \{-1, 1\}, y_i \in \{-1, 1\}$$

$$p(x, y) = \frac{1}{Z} \exp \{-E(x, y)\}$$

$$E(x, y) = h \sum_i x_i - \beta \sum_{i,j} x_i x_j - \eta \sum_i x_i y_i$$

Koşullu bağımsızlık

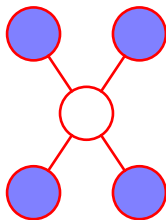


A 'dan B 'ye bütün yollar C 'den geçiyorsa,

$$A \perp B | C$$

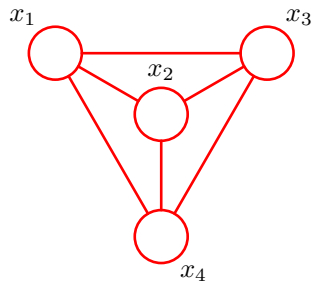
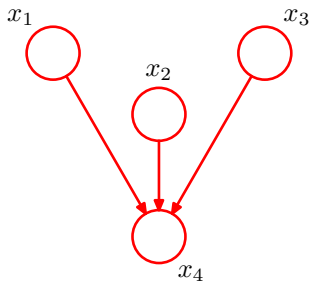
Markov battaniyesi

Bir kenar yakınlığındaki bütün düğümler.



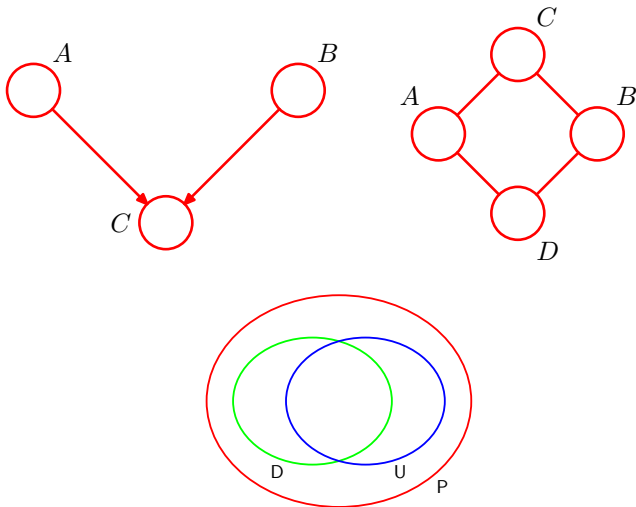
Bir düğüm, Markov battaniyesi verildiğinde diğer düğümlerden bağımsızdır.

Yönlü düz çizgilerden yönsüz çizgelere



- Okların yönlerini sil.
- İlişkileri “ahlakileştir”. (Çocuk ve ebeveynleri arasında klik oluştur.)

Hangisi koşullu bağımsızlığı ifade etmek için daha yetkin?



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 π olan bir Markov zincirinin tasarımına dayanır.

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

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

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

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,

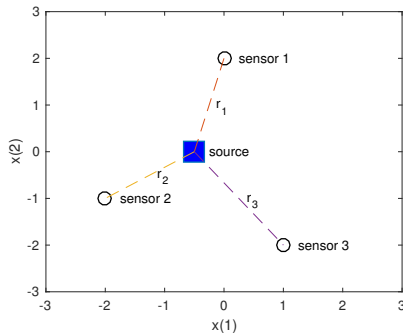
- ▶ Yeni değer için $x' \sim q(\cdot|x)$ önerilir.
- ▶ X_n 'in değeri

$$\alpha(x, x') = \min \left\{ 1, \frac{\pi(x')q(x|x')}{\pi(x)q(x'|x)} \right\}$$

olasılıkla $X_n = x'$ alınır, yoksa önerilen değer reddedilir ve $X_n = x$ alınır.

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:

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

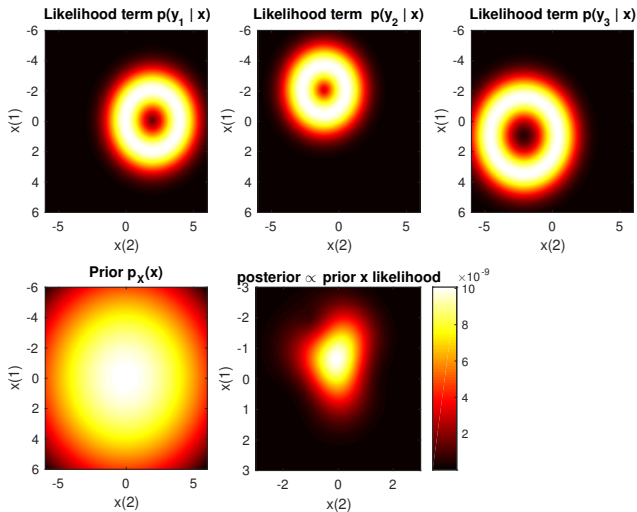
Amaç: $Y = y = (y_1, y_2, y_3)$ verildiğinde $p(x|y)$ 'i bulmak.

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

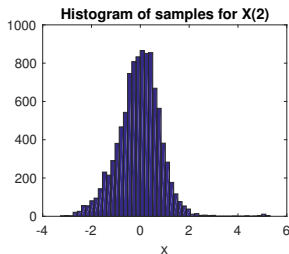
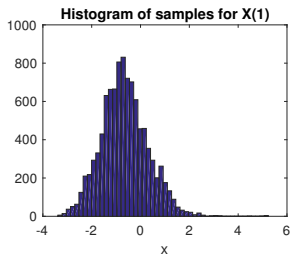
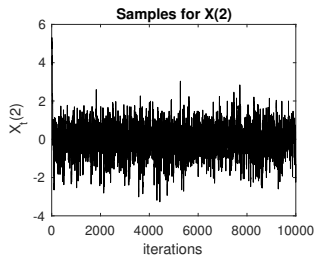
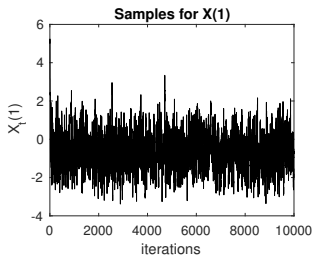
Metropolis-Hastings $q(x'|x) = \phi(x'; x, \sigma_q^2 I_2)$ alınarak çalıştırılabilir:

$$\alpha(x, x') = \min \left\{ 1, \frac{p(x')p(y|x')q(x|x')}{p(x)p(y|x)q(x'|x)} \right\} = \min \left\{ 1, \frac{p(x')p(y|x')}{p(x)p(y|x)} \right\}$$

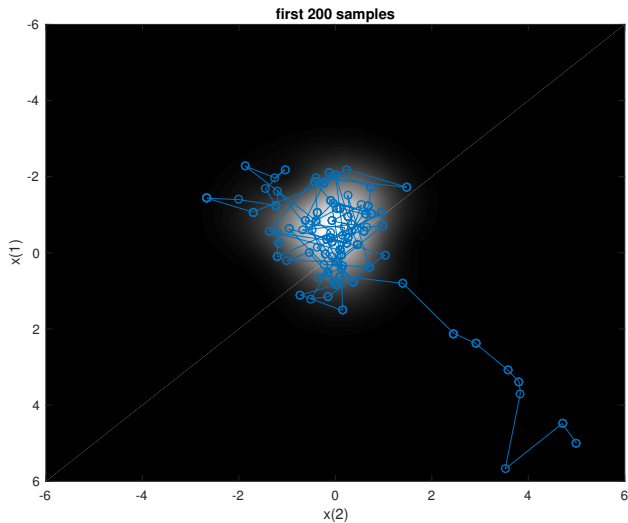
Hedef yer saptaması - dağılımlar, $\sigma_x^2 = 100$, $\sigma_y^2 = 1$



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.

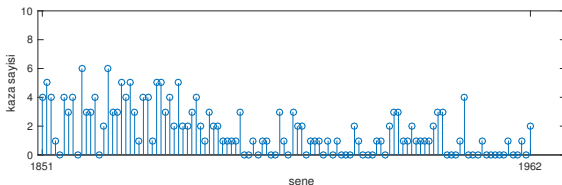
$n = 2, 3, \dots$ için,

$k = 1, \dots, d$ için

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

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:

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

Değişim noktası modelleri - Gibbs örnekleyicisi

Bileşik dağılım fonksiyonu:

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

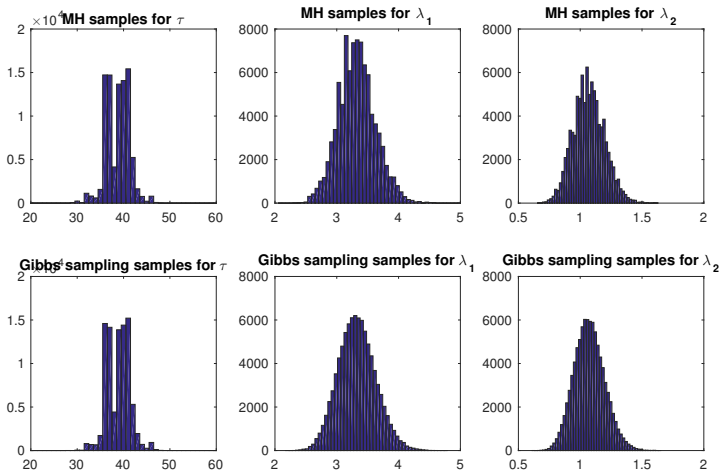
Koşullu dağılımlar:

$$\begin{aligned} \lambda_1 | \tau, \lambda_2, y_{1:n} &\sim \text{Gamma} \left(\alpha + \sum_{t=1}^{\tau} y_t, \beta + \tau \right) \\ \lambda_2 | \tau, \lambda_1, y_{1:n} &\sim \text{Gamma} \left(\alpha + \sum_{t=\tau+1}^n y_t, \beta + n - \tau \right) \\ \tau | \lambda_1, \lambda_2, y_{1:n} &\sim \text{Categorical}(a_1, \dots, a_n) \end{aligned}$$

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

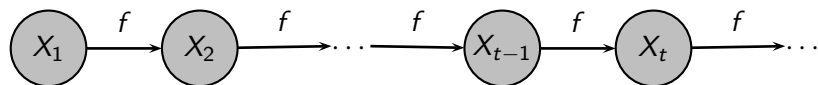
Değişim noktası modelleri - sonuçlar

MH ve Gibbs örnekleme 100000 döngü boyunca çalıştırıldı.



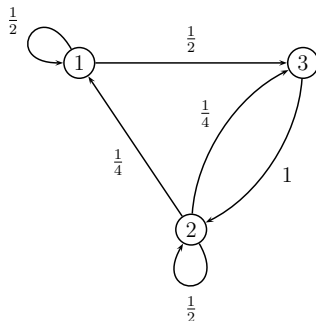
Saklı Markov modelleri

Markov zinciri



$$p(x_1, x_2, \dots, x_n) = f(x_1)f(x_2|x_1) \dots f(x_n|x_{n-1})$$

Markov zinciri - örnek

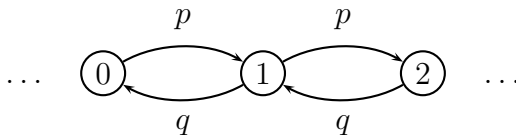


Geçiş olasılıkları matrisi: $F_{i,j} = P(X_{n+1} = j | X_n = i)$

$$F = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{bmatrix}$$

Markov zinciri - örnek

Tamsayılar üzerinde rastgele yürüyüş



$$X_n = X_{n-1} + V_n,$$

$$V_n \in \{-1, 1\}, \quad \mathbb{P}(V_n = 1) = p = 1 - q.$$

Reel sayılar kümesi üzerinde Markov zinciri

$$X_t \in \mathbb{R}.$$

Örnek:

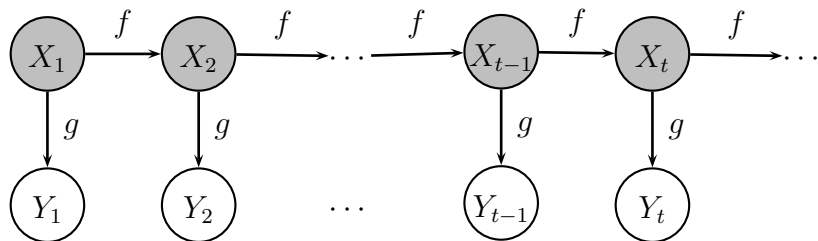
- ▶ Reel sayılar üzerinde rassal yürüyüş:

$$X_n = X_{n-1} + V_n, \quad V_n \sim \mathcal{N}(0, \sigma^2)$$

- ▶ AR(1):

$$X_n = aX_{n-1} + V_n, \quad V_n \sim \mathcal{N}(0, \sigma^2)$$

Saklı Markov modeli



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

- ▶ Saklı süreç: X_1, X_2, \dots
- ▶ Gözlemlenen süreç: Y_1, Y_2, \dots

$$X_t \in \{1, \dots, k\}.$$

Istenecek olasılıklar:

- Süzgeçleme

$$\alpha_t(i) = \mathbb{P}(X_t = i | Y_{1:t} = y_{1:t})$$

- Tahmin

$$\beta_t(i) = \mathbb{P}(X_t = i | Y_{1:t-1} = y_{1:t-1})$$

- Düzleştirme

$$\gamma_t(i) = \mathbb{P}(X_t = i | Y_{1:n} = y_{1:n})$$

İleri-geri algoritması

İleriye doğru süzgeçleme

1 **for** $t = 1, \dots, n$ **do**

2 **Tahmin:** Eğer $t = 1$ ise, $\beta_1(i) = \eta(i)$, $i = 1, \dots, k$; değilse

$$\beta_t(i) = \sum_{j=1}^k \alpha_{t-1}(j) f(i|j), \quad i = 1, \dots, k.$$

Süzgeçleme:

$$\alpha_t(i) = \frac{\beta_t(i) g(y_t|i)}{\sum_{j=1}^k \beta_t(j) g(y_t|j)}, \quad i = 1, \dots, k.$$

Geriye doğru düzleştirme

3 **for** $t = n, \dots, 1$ **do**

4 **Düzleştirme:** Eğer $t = n$ ise, $\gamma_n(i) = \alpha_n(i)$, $i = 1, \dots, k$; değilse

$$\gamma_t(i) = \sum_{j=1}^k \gamma_{t+1}(j) \frac{\alpha_t(i) f(j|i)}{\beta_{t+1}(j)}, \quad i = 1, \dots, k.$$

Doğrusal ve Gauss dinamikli SMM

$$X_1 \sim \mathcal{N}(\mu_1, \Sigma_1), \quad X_t = AX_{t-1} + U_t, \quad U_t \sim \mathcal{N}(0, S), \quad t > 1$$
$$Y_t = BX_t + V_t, \quad V_t \sim \mathcal{N}(0, R)$$

İstenen dağılımlar:

► Süzgeçleme

$$p(x_t | y_{1:t}) = \mathcal{N}(x_t; \mu_{t|t}, P_{t|t}),$$

► Tahmin

$$p(x_t | y_{1:t-1}) = \mathcal{N}(x_t; \mu_{t|t-1}, P_{t|t-1}),$$

► Düzleştirme

$$p(x_t | y_{1:n}) = \mathcal{N}(x_t; \mu_{t|n}, P_{t|n}).$$

İleri-geri algoritması

Kalman süzgeci (İleriye doğru süzgeçleme)

1 for $t = 1, \dots, n$ do

2 Tahmin: If $t = 1$, $\mu_{1|0} = \mu_1$, $P_{1|0} = \Sigma_1$; değilse

$$\mu_{t|t-1} = A\mu_{t-1|t-1}$$

$$P_{t|t-1} = AP_{t-1|t-1}A^T + S$$

Süzgeçleme:

$$P_{t|t-1}^y = BP_{t|t-1}B^T + R$$

$$\mu_{t|t-1}^y = B\mu_{t|t-1}$$

$$P_{t|t-1}^{xy} = P_{t|t-1}B^T$$

$$\mu_{t|t} = \mu_{t|t-1} + P_{t|t-1}^{xy} P_{t|t-1}^{yy}{}^{-1} (y_t - \mu_{t|t-1}^y)$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}^{xy} P_{t|t-1}^{yy}{}^{-1} P_{t|t-1}^{xyT}$$

Geriye doğru düzeltirme

3 for $t = n - 1, \dots, 1$ do

4 Düzleştirme:

$$\Gamma_{t|t+1} = P_{t|t}A^T P_{t+1|t}^{-1}$$

$$\mu_{t|n} = \mu_{t|t} + \Gamma_{t|t+1}(\mu_{t+1|n} - \mu_{t+1|t})$$

$$P_{t|n} = P_{t|t} + \Gamma_{t|t+1}(P_{t+1|n} - P_{t+1|t})\Gamma_{t|t+1}^T$$

Son söz yerine: Üç nasihat



1. **Olmaz deme:** Bir değişkenle ilgili öncül dağılımı seçerken, doğal sınırları içinde alabileceği bir değere 0 ihtimal verme.
2. **Ön yargılı ol:** Yeni veriyi işlerken önceden bildiklerini kullan, **ama**
3. **Sabit fikirli olma:** Ön yargını yeni bir tecrübeden (veriden) sonra güncelleyerek son yargıya dönüştür.