

Ostatni wykład z Algorytmicznej Teorii Gier

Dominik Bojko, Politechnika Wrocławska

1 lutego 2024

Total Functions Nondeterministic Polynomial (TFNP)

Relacja $R(x, y)$ jest w TFNP IFF

- 1 Istnieje deterministyczny alg. o złożoności najwyżej wielomianowej, który może rozstrzygnąć, czy $R(x, y)$,
- 2 $(\forall x)(\exists y)$ y jest „wielkością” co najwyżej wielomianej względem x oraz spełniona jest relacja $R(x, y)$.

Total Functions Nondeterministic Polynomial (TFNP)

Relacja $R(x, y)$ jest w TFNP IFF

- 1 Istnieje deterministyczny alg. o złożoności najwyżej wielomianowej, który może rozstrzygnąć, czy $R(x, y)$,
- 2 $(\forall x)(\exists y)$ y jest „wielkością” co najwyżej wielomianej względem x oraz spełniona jest relacja $R(x, y)$.

Przykład: Zasada szufladkowa Dirichleta

Dana jest funkcja $f : [n + 1] \rightarrow [n]$, która jest wielomianowo obliczalna. Znajdź $a, b \in [n + 1]$, że $f(a) = f(b)$.

Wtedy $x = f$, a $y = \{k, l\} \subset [n + 1]$, a $R(f, \{k, l\})$ oznacza, że $f[\{k\}] = f[\{l\}]$.

Total Functions Nondeterministic Polynomial (TFNP)

Relacja $R(x, y)$ jest w TFNP IFF

- 1 Istnieje deterministyczny alg. o złożoności najwyżej wielomianowej, który może rozstrzygnąć, czy $R(x, y)$,
- 2 $(\forall x)(\exists y)$ y jest „wielkością” co najwyżej wielomianej względem x oraz spełniona jest relacja $R(x, y)$.

Przykład: Zasada szufladkowa Dirichleta

Dana jest funkcja $f : [n + 1] \rightarrow [n]$, która jest wielomianowo obliczalna. Znajdź $a, b \in [n + 1]$, że $f(a) = f(b)$.

Wtedy $x = f$, a $y = \{k, l\} \subset [n + 1]$, a $R(f, \{k, l\})$ oznacza, że $f[\{k\}] = f[\{l\}]$.

$\text{TFNP} \subseteq \text{NP}$

Polynomial Parity Arguments on Directed graphs (PPAD)

Problem PPAD to taki, który można sprowadzić do problemu: End-of-the-Line.

Polynomial Parity Arguments on Directed graphs (PPAD)

Problem PPAD to taki, który można sprowadzić do problemu:
End-of-the-Line.

Dany jest graf skierowany G (może być nawet eksponencjalnie duży), gdzie każdy wierzchołek może mieć co najwyżej 1 poprzednika i 1 następnika.

Polynomial Parity Arguments on Directed graphs (PPAD)

Problem PPAD to taki, który można sprowadzić do problemu:
End-of-the-Line.

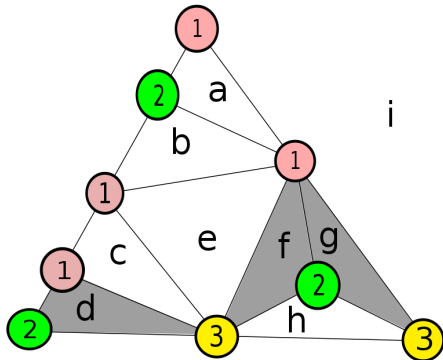
Dany jest graf skierowany G (może być nawet eksponencjalnie duży), gdzie każdy wierzchołek może mieć co najwyżej 1 poprzednika i 1 następnika.

Zadana jest funkcja $f : V \rightarrow V' \times V'$, o złożoności co najwyżej wielomianowej względem $|V|$, gdzie $V' = V \cup \{\emptyset\}$, która danemu wierzchołkowi, przypisuje jego poprzednika oraz następnika (o ile istnieją).

Problem: Dla $s \in V$, który nie ma poprzednika, szukamy $t \neq s$, który nie ma poprzednika albo następnika.

Przykłady problemów PPAD

- Lemat Spenera



Przykłady problemów PPAD

- Lemat Spernera
- Lemat Knastera—Kuratowskiego—Mazurkiewicza ($\sim$ spójność sympleksu)

Przykłady problemów PPAD

- Lemat Spernera
- Lemat Knastera—Kuratowskiego—Mazurkiewicza ($\sim$ spójność sympleksu)
- Szukanie punktów stałych w tw. Brouwera.

Przykłady problemów PPAD

- Lemat Spernera
- Lemat Knastera—Kuratowskiego—Mazurkiewicza ($\sim$ spójność sympleksu)
- Szukanie punktów stałych w tw. Brouwera.
- Szukanie MNE
- Szukanie ogólnej równowagi (w tzw. modelu Arrowa—Debrou [1954; nagrody Nobla 1972 i 1983]).

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa (d-d lem. Spernera/ tw. Kakutaniego).

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa (d-d lem. Spernera/ tw. Kakutaniego).

Definicja (Równowaga przybliżona)

Dla ustalonego $\varepsilon > 0$, dla gry (P, S, u) , $\pi^ \in \Delta$ nazywamy ε -równowagą Nasha, jeśli spełnia:*

$$(\forall i \in P)(\forall \pi_i \in \Delta_i) u_i(\pi^*) \geq u_i(\langle \pi_i, \pi_{-i}^* \rangle) - \varepsilon .$$

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa (d-d lem. Spernera/ tw. Kakutaniego).

Definicja (Równowaga przybliżona)

Dla ustalonego $\varepsilon > 0$, dla gry (P, S, u) , $\pi^ \in \Delta$ nazywamy ε -równowagą Nasha, jeśli spełnia:*

$$(\forall i \in P)(\forall \pi_i \in \Delta_i) u_i(\pi^*) \geq u_i(\langle \pi_i, \pi_{-i}^* \rangle) - \varepsilon .$$

Alg. wyznaczania równowag przybliżonych są szybsze od tych dla MNE, np. metoda Newtona.

Ale mogą być dalekie od MNE.

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa (d-d lem. Spernera/ tw. Kakutaniego).

Definicja (Równowaga przybliżona)

Dla ustalonego $\varepsilon > 0$, dla gry (P, S, u) , $\pi^ \in \Delta$ nazywamy ε -równowagą Nasha, jeśli spełnia:*

$$(\forall i \in P)(\forall \pi_i \in \Delta_i) u_i(\pi^*) \geq u_i(\langle \pi_i, \pi_{-i}^* \rangle) - \varepsilon .$$

Alg. wyznaczania równowag przybliżonych są szybsze od tych dla MNE, np. metoda Newtona.

Ale mogą być dalekie od MNE.

Algorytm Scarfa nadaje się do szukania „lepszyc” równowag przybliżonych.

Twierdzenie (Kuhn)

W lematkach Spernera oraz KKM można zamienić sympleksy na hiperkostki.

Szukanie równowag Nasha w grach strategicznych wieloosobowych — alg. Scarfa (d-d lem. Spernera/ tw. Kakutaniego).

Definicja (Równowaga przybliżona)

Dla ustalonego $\varepsilon > 0$, dla gry (P, S, u) , $\pi^ \in \Delta$ nazywamy ε -równowagą Nasha, jeśli spełnia:*

$$(\forall i \in P)(\forall \pi_i \in \Delta_i) u_i(\pi^*) \geq u_i(\langle \pi_i, \pi_{-i}^* \rangle) - \varepsilon .$$

Alg. wyznaczania równowag przybliżonych są szybsze od tych dla MNE, np. metoda Newtona.

Ale mogą być dalekie od MNE.

Algorytm Scarfa nadaje się do szukania „lepszyc” równowag przybliżonych.

Problem klasy PPAD.

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .
Jaką cenę powinni podyktować?

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .

Jaką cenę powinni podyktować?

- Jeśli zagrają $v_i \leq p$, to nie zyskają.

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .

Jaką cenę powinni podyktować?

- Jeśli zagrają $v_i \leq p$, to nie zyskają.
- Jeśli zagrają $v_i > v_j > p$, to wszyscy klienci kupią produkt po cenie v_j i firma j zgarnie cały zysk.

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .

Jaką cenę powinni podyktować?

- Jeśli zagrają $v_i \leq p$, to nie zyskają.
- Jeśli zagrają $v_i > v_j > p$, to wszyscy klienci kupią produkt po cenie v_j i firma j zgarnie cały zysk.
- Jeśli $v_i = v_j > p$, to firmy podzielą się zyskiem po połowie.

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .

Jaką cenę powinni podyktować?

- Jeśli zagrają $v_i \leq p$, to nie zyskają.
- Jeśli zagrają $v_i > v_j > p$, to wszyscy klienci kupią produkt po cenie v_j i firma j zgarnie cały zysk.
- Jeśli $v_i = v_j > p$, to firmy podzielą się zyskiem po połowie.
- W tej grze jedyna równowaga skorelowana to zagrać $v_i = v_j = p$.

Duopol Bertranda

Dokładnie dwie firmy produkują to samo dobro; w dodatku po tym samym marginalnym koszcie p .

Jaką cenę powinni podyktować?

- Jeśli zagrają $v_i \leq p$, to nie zyskają.
- Jeśli zagrają $v_i > v_j > p$, to wszyscy klienci kupią produkt po cenie v_j i firma j zgarnie cały zysk.
- Jeśli $v_i = v_j > p$, to firmy podzielą się zyskiem po połowie.
- W tej grze jedyna równowaga skorelowana to zagrać $v_i = v_j = p$.
- Zysk może zagwarantować im tutaj koalicja (zmowa cenowa).

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,
- Firmy mają siłę rynkową, czyli zmiana decyzji firmy na wpływ na rynek

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,
- Firmy mają siłę rynkową, czyli zmiana decyzji firmy ma wpływ na rynek
- Sprzedają do tego samego pośrednika $\Rightarrow$ konkurencja na polu ilościowym,

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,
- Firmy mają siłę rynkową, czyli zmiana decyzji firmy ma wpływ na rynek
- Sprzedają do tego samego pośrednika $\Rightarrow$ konkurencja na polu ilościowym,
- Firmy maksymalizują ich profit, znając decyzje innych.

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,
- Firmy mają siłę rynkową, czyli zmiana decyzji firmy ma wpływ na rynek
- Sprzedają do tego samego pośrednika $\Rightarrow$ konkurencja na polu ilościowym,
- Firmy maksymalizują ich profit, znając decyzje innych.
- Zapotrzebowanie rynku D jest malejącą funkcją ceny p : $D = c(p)$.

Oligopol Cournota

Założenia modelu:

- Kilka firm (ustalonych) produkuje ten sam produkt po tym samym koszcie,
- Nie ma zmów cenowych,
- Firmy mają siłę rynkową, czyli zmiana decyzji firmy ma wpływ na rynek
- Sprzedają do tego samego pośrednika $\Rightarrow$ konkurencja na polu ilościowym,
- Firmy maksymalizują ich profit, znając decyzje innych.
- Zapotrzebowanie rynku D jest malejącą funkcją ceny p : $D = c(p)$.
- Zysk ze sprzedaży to $pc(p)$. Maksymalizacja tej funkcji wiąże się z równaniem $c(p) + pc'(p) = 0$.

Przykład

- Dwie firmy, z kosztem marginalnym p

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapostrzebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapożrebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.
- Strategie firm to produkcja odpowiednio: q_1, q_2 dobra.

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapożrebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.
- Strategie firm to produkcja odpowiednio: q_1, q_2 dobra.
- Zysk: $u_i = (a - bq_1 - bq_2 - p)q_i$.

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapożrebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.
- Strategie firm to produkcja odpowiednio: q_1, q_2 dobra.
- Zysk: $u_i = (a - bq_1 - bq_2 - p)q_i$.
- $u'_i = a - 2bq_i - bq_{2-i} - p = 0$

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapożrebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.
- Strategie firm to produkcja odpowiednio: q_1, q_2 dobra.
- Zysk: $u_i = (a - bq_1 - bq_2 - p)q_i$.
- $u'_i = a - 2bq_i - bq_{2-i} - p = 0$
- Równowaga Cournota—Nasha to $q_i = \frac{a-p-bq_{2-i}}{2b}$

Przykład

- Dwie firmy, z kosztem marginalnym p
- Cena, jako funkcja zapożrzebowania D : $f(D) = c^{-1}(D) = a - bD$, gdzie $a > p$, $b > 0$.
- Strategie firm to produkcja odpowiednio: q_1, q_2 dobra.
- Zysk: $u_i = (a - bq_1 - bq_2 - p)q_i$.
- $u'_i = a - 2bq_i - bq_{2-i} - p = 0$
- Równowaga Cournota—Nasha to $q_i = \frac{a-p-bq_{2-i}}{2b}$
- Z symetrii: $q_1 = q_2 = \frac{a-p}{3b}$.

Model Stackelberga

Leader-follower model — gra w postaci ekstensywnej,
(firmy wykonują ruchy od najlepszej, do najgorszej)

Model Stackelberga

Leader-follower model — gra w postaci ekstensywnej,
(firmy wykonują ruchy od najlepszej, do najgorszej)

Połączenie:

- EFG,
- modelu gier oligopolowych,
- gier bayesowskich (lider ma rozkład a priori ruchów pozostałych)

Model Stackelberga

Leader-follower model — gra w postaci ekstensywnej,
(firmy wykonują ruchy od najlepszej, do najgorszej)

Połączenie:

- EFG,
- modelu gier oligopolowych,
- gier bayesowskich (lider ma rozkład a priori ruchów pozostałych)
- rozważa się też oligopolowe gry na sieciach (firmy sprzedają przesył przez krawędzie)

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej:
 $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej: $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

- Θ — *strategie natury (odp. S_0)*
- $\mathcal{A}$ — *strategie statystyka (odp. S_1)*

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej: $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

- Θ — *strategie natury (odp. S_0)*
- $\mathcal{A}$ — *strategie statystyka (odp. S_1)*
- $(\mathcal{X}, \mathcal{B})$ — *przestrzeń mierzalna*

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej: $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

- Θ — *strategie natury (odp. S_0)*
- $\mathcal{A}$ — *strategie statystyka (odp. S_1)*
- $(\mathcal{X}, \mathcal{B})$ — *przestrzeń mierzalna*
- $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ — *rodzina dopuszczalnych rozkładów na $(\mathcal{X}, \mathcal{B})$*
Każda $(\mathcal{X}, \mathcal{B}, P_\theta)$ — przestrzeń probabilistyczna
 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ — przestrzeń statystyczna

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej:
 $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

- Θ — *strategie natury (odp. S_0)*
- $\mathcal{A}$ — *strategie statystyka (odp. S_1)*
- $(\mathcal{X}, \mathcal{B})$ — *przestrzeń mierzalna*
- $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ — *rodzina dopuszczalnych rozkładów na $(\mathcal{X}, \mathcal{B})$*
Każda $(\mathcal{X}, \mathcal{B}, P_\theta)$ — przestrzeń probabilistyczna
 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ — przestrzeń statystyczna
- $\mathcal{L} : \Theta \times \mathcal{A} \rightarrow \mathbb{R}$ — *funkcja straty (odp. u)*

Definicja

Gra statystyczna — gra jednoosobowa w postaci normalnej: $(\{0, 1\}, \Theta \times \mathcal{A}, -\mathcal{L})$, gdzie:

- Θ — *strategie natury (odp. S_0)*
- $\mathcal{A}$ — *strategie statystyka (odp. S_1)*
- $(\mathcal{X}, \mathcal{B})$ — *przestrzeń mierzalna*
- $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ — *rodzina dopuszczalnych rozkładów na $(\mathcal{X}, \mathcal{B})$*
Każda $(\mathcal{X}, \mathcal{B}, P_\theta)$ — przestrzeń probabilistyczna
 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ — przestrzeń statystyczna
- $\mathcal{L} : \Theta \times \mathcal{A} \rightarrow \mathbb{R}$ — *funkcja straty (odp. u)*

Natura losuje coś z pewnego rozkładu P_θ , a statystyk akcję $a \in \mathcal{A}$.

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

[Realizacja $X \sim P_\theta$ to x] $\Rightarrow$ [$d(x)$ to podjęta akcja gracza].

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

[Realizacja $X \sim P_\theta$ to x] $\Rightarrow$ [$d(x)$ to podjęta akcja gracza].

$\mathcal{L}(\theta, d(x))$ — strata gracza przy tej realizacji gry.

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

[Realizacja $X \sim P_\theta$ to x] $\Rightarrow$ [$d(x)$ to podjęta akcja gracza].

$\mathcal{L}(\theta, d(x))$ — strata gracza przy tej realizacji gry.

Definicja

Funkcję ryzyka $\mathcal{R} : \Theta \times \mathcal{D} \rightarrow \mathbb{R}$ określamy jako:

$$\mathcal{R}(\theta, d) = \mathbb{E}_\theta[\mathcal{L}(\theta, d(X))] = \int_{\mathcal{X}} \mathcal{L}(\theta, d(x)) dP_\theta(x) .$$

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

[Realizacja $X \sim P_\theta$ to x] $\Rightarrow$ [$d(x)$ to podjęta akcja gracza].

$\mathcal{L}(\theta, d(x))$ — strata gracza przy tej realizacji gry.

Definicja

Funkcję ryzyka $\mathcal{R} : \Theta \times \mathcal{D} \rightarrow \mathbb{R}$ określamy jako:

$$\mathcal{R}(\theta, d) = \mathbb{E}_\theta[\mathcal{L}(\theta, d(X))] = \int_{\mathcal{X}} \mathcal{L}(\theta, d(x)) dP_\theta(x).$$

Odp. $-u(\pi)$, gdzie $\pi = (\pi_0, \pi_1) = (P_\theta, d)$,
gdy $\mathcal{P} = \Delta_0$ i $S(\pi_1) = \{d\} \subset \mathcal{D} = S_1$.

Funkcje decyzyjne i ryzyka

Definicja

Deterministyczna funkcja decyzyjna $d : \mathcal{X} \rightarrow \mathcal{A}$ — f. mierzalna, gdzie $(\mathcal{X}, \mathcal{B}), (\mathcal{A}, \mathcal{G})$ — pp. mierzalne.

$\mathcal{D}$ — zb. wszystkich deterministycznych funkcji decyzyjnych.

[Realizacja $X \sim P_\theta$ to x] $\Rightarrow$ [$d(x)$ to podjęta akcja gracza].

$\mathcal{L}(\theta, d(x))$ — strata gracza przy tej realizacji gry.

Definicja

Funkcję ryzyka $\mathcal{R} : \Theta \times \mathcal{D} \rightarrow \mathbb{R}$ określamy jako:

$$\mathcal{R}(\theta, d) = \mathbb{E}_\theta[\mathcal{L}(\theta, d(X))] = \int_{\mathcal{X}} \mathcal{L}(\theta, d(x)) dP_\theta(x).$$

Odp. $-u(\pi)$, gdzie $\pi = (\pi_0, \pi_1) = (P_\theta, d)$,

gdy $\mathcal{P} = \Delta_0$ i $S(\pi_1) = \{d\} \subset \mathcal{D} = S_1$.

Gra decyzyjna: $(\{0, 1\}, \Theta \times \mathcal{D}, -\mathcal{R})$

Randomizacja funkcji decyzyjnych I

Definicja

Niech $(\mathcal{D}, \mathcal{F}, \delta^)$ -p. probabilistyczna, gdzie $\{\{d\} : d \in \mathcal{D}\} \subseteq \mathcal{F}$.*

Randomizacja funkcji decyzyjnych I

Definicja

Niech $(\mathcal{D}, \mathcal{F}, \delta^)$ -p. probabilistyczna, gdzie $\{\{d\} : d \in \mathcal{D}\} \subseteq \mathcal{F}$. δ^* nazywamy mieszaną funkcją decyzyjną.*

$\mathcal{D}^$ —zb. mieszanych funkcji decyzyjnych.*

Randomizacja funkcji decyzyjnych I

Definicja

Niech $(\mathcal{D}, \mathcal{F}, \delta^*)$ -p. probabilistyczna, gdzie $\{\{d\} : d \in \mathcal{D}\} \subseteq \mathcal{F}$. δ^* nazywamy mieszaną funkcją decyzyjną.

$\mathcal{D}^*$ —zb. mieszanych funkcji decyzyjnych.

$$\mathcal{R}^*(\theta, \delta^*) = \mathbb{E}^{\delta^*} \mathcal{R}(\theta, D) = \int_{\mathcal{D}} \mathcal{R}(\theta, \tilde{d}) d\delta^*(\tilde{d})$$

Randomizacja funkcji decyzyjnych I

Definicja

Niech $(\mathcal{D}, \mathcal{F}, \delta^*)$ -p. probabilistyczna, gdzie $\{\{d\} : d \in \mathcal{D}\} \subseteq \mathcal{F}$. δ^* nazywamy mieszaną funkcją decyzyjną.

$\mathcal{D}^*$ —zb. mieszanych funkcji decyzyjnych.

$$\mathcal{R}^*(\theta, \delta^*) = \mathbb{E}^{\delta^*} \mathcal{R}(\theta, D) = \int_{\mathcal{D}} \mathcal{R}(\theta, \tilde{d}) d\delta^*(\tilde{d})$$

Odp. $u(\pi)$

Randomizacja funkcji decyzyjnych I

Definicja

Niech $(\mathcal{D}, \mathcal{F}, \delta^*)$ -p. probabilistyczna, gdzie $\{\{d\} : d \in \mathcal{D}\} \subseteq \mathcal{F}$. δ^* nazywamy mieszaną funkcją decyzyjną.

$\mathcal{D}^*$ —zb. mieszanych funkcji decyzyjnych.

$$\mathcal{R}^*(\theta, \delta^*) = \mathbb{E}^{\delta^*} \mathcal{R}(\theta, D) = \int_{\mathcal{D}} \mathcal{R}(\theta, \tilde{d}) d\delta^*(\tilde{d})$$

Odp. $u(\pi)$

Zrandomizowana gra decyzyjna: $(\{0, 1\}, \Theta \times \mathcal{D}^*, -\mathcal{R}^*)$.

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, \mathcal{Q})$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, Q)$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

$$\bar{\mathcal{L}}(\theta, Q) = \mathbb{E}^Q[\mathcal{L}(\theta, A)] = \int_{\mathcal{A}} \mathcal{L}(\theta, a) dQ(a)$$

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, \mathcal{Q})$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

$$\bar{\mathcal{L}}(\theta, Q) = \mathbb{E}^Q[\mathcal{L}(\theta, A)] = \int_{\mathcal{A}} \mathcal{L}(\theta, a) dQ(a)$$

$\bar{\delta} : \mathcal{X} \times \mathcal{G} \rightarrow \mathcal{Q}$ — to behawiorystyczna f. decyzyjną, jeśli:

- $(\forall G \in \mathcal{G}) \bar{\delta}(\cdot, G)$ jest $\mathcal{B}$ -mierzalna,
- $(\forall x \in \mathcal{X}) \bar{\delta}(x, \cdot)$ jest miarą probabilistyczną na $(\mathcal{A}, \mathcal{G})$.

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, \mathcal{Q})$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

$$\bar{\mathcal{L}}(\theta, Q) = \mathbb{E}^Q[\mathcal{L}(\theta, A)] = \int_{\mathcal{A}} \mathcal{L}(\theta, a) dQ(a)$$

$\bar{\delta} : \mathcal{X} \times \mathcal{G} \rightarrow \mathcal{Q}$ — to behawiorystyczna f. decyzyjną, jeśli:

- $(\forall G \in \mathcal{G}) \bar{\delta}(\cdot, G)$ jest $\mathcal{B}$ -mierzalna,
- $(\forall x \in \mathcal{X}) \bar{\delta}(x, \cdot)$ jest miarą probabilistyczną na $(\mathcal{A}, \mathcal{G})$.

$\bar{\mathcal{D}}$ — zb. behawiorystycznych funkcji decyzyjnych.

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, \mathcal{Q})$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

$$\bar{\mathcal{L}}(\theta, Q) = \mathbb{E}^Q[\mathcal{L}(\theta, A)] = \int_{\mathcal{A}} \mathcal{L}(\theta, a) dQ(a)$$

$\bar{\delta} : \mathcal{X} \times \mathcal{G} \rightarrow \mathcal{Q}$ — to behawiorystyczna f. decyzyjna, jeśli:

- $(\forall G \in \mathcal{G}) \bar{\delta}(\cdot, G)$ jest $\mathcal{B}$ -mierzalna,
- $(\forall x \in \mathcal{X}) \bar{\delta}(x, \cdot)$ jest miarą probabilistyczną na $(\mathcal{A}, \mathcal{G})$.

$\bar{\mathcal{D}}$ — zb. behawiorystycznych funkcji decyzyjnych.

$$\begin{aligned} \bar{\mathcal{R}}(\theta, \bar{\delta}) &= \mathbb{E}_{\theta} \bar{\mathcal{L}}(\theta, \bar{\delta}(X, \cdot)) = \mathbb{E}_{\theta} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta, A) \right] = \\ &= \int \int_{\mathcal{X} \times \mathcal{A}} \mathcal{L}(\theta, a) d\bar{\delta}(a) dP_{\theta}(x) \end{aligned}$$

Randomizacja funkcji decyzyjnych II

Definicja

Niech $(\mathcal{A}, \mathcal{G}, \mathcal{Q})$ — p. statystyczna, gdzie $\{\{a\} : a \in \mathcal{A}\} \subseteq \mathcal{G}$.

$$\bar{\mathcal{L}}(\theta, \mathcal{Q}) = \mathbb{E}^{\mathcal{Q}}[\mathcal{L}(\theta, A)] = \int_{\mathcal{A}} \mathcal{L}(\theta, a) d\mathcal{Q}(a)$$

$\bar{\delta} : \mathcal{X} \times \mathcal{G} \rightarrow \mathcal{Q}$ — to behawiorystyczna f. decyzyjna, jeśli:

- $(\forall G \in \mathcal{G}) \bar{\delta}(\cdot, G)$ jest $\mathcal{B}$ -mierzalna,
- $(\forall x \in \mathcal{X}) \bar{\delta}(x, \cdot)$ jest miarą probabilistyczną na $(\mathcal{A}, \mathcal{G})$.

$\bar{\mathcal{D}}$ — zb. behawiorystycznych funkcji decyzyjnych.

$$\begin{aligned} \bar{\mathcal{R}}(\theta, \bar{\delta}) &= \mathbb{E}_{\theta} \bar{\mathcal{L}}(\theta, \bar{\delta}(X, \cdot)) = \mathbb{E}_{\theta} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta, A) \right] = \\ &= \int \int_{\mathcal{X} \times \mathcal{A}} \mathcal{L}(\theta, a) d\bar{\delta}(a) dP_{\theta}(x) \end{aligned}$$

Behawiorystyczna gra decyzyjna: $(\{0, 1\}, \Theta \times \bar{\mathcal{D}}, -\bar{\mathcal{R}})$.

Definicja

Niech (X, ρ) — p. metryczna. Y nazywamy ośrodkiem w (X, ρ) , jeśli

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists y \in Y) \rho(x, y) < \varepsilon$$

Definicja

Niech (X, ρ) — p. metryczna. Y nazywamy ośrodkiem w (X, ρ) , jeśli

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists y \in Y) \rho(x, y) < \varepsilon$$

Innymi słowy: Y jest zbiorem gęstym w X .

Definicja

Niech (X, ρ) — p . metryczna. Y nazywamy ośrodkiem w (X, ρ) , jeśli

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists y \in Y) \rho(x, y) < \varepsilon$$

Innymi słowy: Y jest zbiorem gęstym w X .

(X, ρ) nazywamy ośrodkową, jeśli ma co najwyżej przeliczalny ośrodek.

Definicja

Niech (X, ρ) — p. metryczna. Y nazywamy ośrodkiem w (X, ρ) , jeśli

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists y \in Y) \rho(x, y) < \varepsilon$$

Innymi słowy: Y jest zbiorem gęstym w X .

(X, ρ) nazywamy ośrodkową, jeśli ma co najwyżej przeliczalny ośrodek.

Przestrzeń metryczna, zwarta jest ośrodkowa.

Definicja

Niech (X, ρ) — p. metryczna. Y nazywamy ośrodkiem w (X, ρ) , jeśli

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists y \in Y) \rho(x, y) < \varepsilon$$

Innymi słowy: Y jest zbiorem gęstym w X .

(X, ρ) nazywamy ośrodkową, jeśli ma co najwyżej przeliczalny ośrodek.

Przestrzeń metryczna, zwarta jest ośrodkowa.

Twierdzenie (Wald—Wolfowitz)

Jeśli $(\mathcal{A}, \rho)$ jest p. ośrodkową, to $(\{0, 1\}, \Theta \times \mathcal{D}^*, -\mathcal{R}^*)$ i $(\{0, 1\}, \Theta \times \bar{\mathcal{D}}, -\bar{\mathcal{R}})$ są równoważne.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.
 $\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

W szczególności, gdy $c_j \equiv 1$, $\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = 1 - \mathbb{E}_{\theta_i} \bar{\delta}(X, \{\theta_i\})$.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

W szczególności, gdy $c_j \equiv 1$, $\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = 1 - \mathbb{E}_{\theta_i} \bar{\delta}(X, \{\theta_i\})$.

Testowanie hipotez: $k = 2$, $\mathcal{A} = \Theta' = \{\theta \in \Theta_0, \theta \in \Theta_1\}$, gdzie

$\Theta = \Theta_0 \cup \Theta_1$.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

W szczególności, gdy $c_j \equiv 1$, $\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = 1 - \mathbb{E}_{\theta_i} \bar{\delta}(X, \{\theta_i\})$.

Testowanie hipotez: $k = 2$, $\mathcal{A} = \Theta' = \{\theta \in \Theta_0, \theta \in \Theta_1\}$, gdzie

$\Theta = \Theta_0 \cup \Theta_1$.

Wtedy c_0 to kara za False Positive (błąd II rodzaju), a c_1 — za False Negative (błąd I rodzaju).

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

W szczególności, gdy $c_j \equiv 1$, $\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = 1 - \mathbb{E}_{\theta_i} \bar{\delta}(X, \{\theta_i\})$.

Testowanie hipotez: $k = 2$, $\mathcal{A} = \Theta' = \{\theta \in \Theta_0, \theta \in \Theta_1\}$, gdzie

$\Theta = \Theta_0 \cup \Theta_1$.

Wtedy c_0 to kara za False Positive (błąd II rodzaju), a c_1 — za False Negative (błąd I rodzaju).

Moc testu: $\beta(\theta) = P_\theta(d(X) = (\theta \in \Theta_1))$.

$R(\theta, d) = c_1 \beta(\theta) \mathbb{I}[\theta \in \Theta_0] + c_0 (1 - \beta(\theta)) \mathbb{I}[\theta \in \Theta_1]$.

Problem klasyfikacji: $\Theta = \{\theta_1, \dots, \theta_k\} = \mathcal{A}$.

$\mathcal{L}(\theta_i, \theta_j) = c_j \cdot \mathbb{I}[i \neq j]$, gdzie $c_j > 0$, stałe.

$$\bar{\delta}(x, A) = \int_A \mathcal{L}(\theta_i, a) d\bar{\delta}(x, a) = \sum_{j=1}^k c_j \mathbb{I}[\theta_j \in A] \bar{\delta}(x, \theta_j),$$

dla $A \in \mathcal{G}$.

$$\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = \mathbb{E}_{\theta_i} \left[\mathbb{E}^{\bar{\delta}(X, \cdot)} \mathcal{L}(\theta_i, A) \right] = \mathbb{E}_{\theta_i} \sum_{j=1, j \neq i}^k c_j \mathbb{I}[\theta_i \notin \{\theta_j\}] \bar{\delta}(X, \{\theta_j\}).$$

W szczególności, gdy $c_j \equiv 1$, $\bar{\mathcal{R}}(\theta_i, \bar{\delta}) = 1 - \mathbb{E}_{\theta_i} \bar{\delta}(X, \{\theta_i\})$.

Testowanie hipotez: $k = 2$, $\mathcal{A} = \Theta' = \{\theta \in \Theta_0, \theta \in \Theta_1\}$, gdzie

$\Theta = \Theta_0 \cup \Theta_1$.

Wtedy c_0 to kara za False Positive (błąd II rodzaju), a c_1 — za False Negative (błąd I rodzaju).

Moc testu: $\beta(\theta) = P_{\theta}(d(X) = (\theta \in \Theta_1))$.

$R(\theta, d) = c_1 \beta(\theta) \mathbb{I}[\theta \in \Theta_0] + c_0 (1 - \beta(\theta)) \mathbb{I}[\theta \in \Theta_1]$.

Zastosowanie: Ai, Machine Learning

Rozkłady a priori

Definicja

Rozkładem a priori na p. mierzalnej (Θ, Ξ) nazywamy miarę probabilistyczną π parametru θ .

Π — zbiór wszystkich rozkładów a priori.

Rozkłady a priori

Definicja

Rozkładem a priori na p. mierzalnej (Θ, Ξ) nazywamy miarę probabilistyczną π parametru θ .

Π — *zbiór wszystkich rozkładów a priori.*

Można myśleć, że P jest rozkładem prawdopodobieństwa wektora losowego (X, Θ) na $(\mathcal{X} \times \Theta, \mathcal{B} \times \Xi)$, a p jego gęstością wzgl. pewnej miary ν .

$$P((X, \Theta) \in A) = \int_A p(x, \theta) d\nu(x, \theta) .$$

Twierdzenie (Bayes)

Gdy $P(B) > 0$, to $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$.

Twierdzenie (Bayes)

Gdy $P(B) > 0$, to $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$.

Twierdzenie (bayesowskie)

Przypuśćmy, że $\pi \in \Pi$ ma gęstość $h(\theta)$ względem miary μ na (Θ, Ξ) .
Dla każdego $x \in X$, istnieje rozkład a posteriori:

$$h(\theta|x) = \frac{p_\theta(x)h(\theta)}{m(x)} = \frac{p(x, \theta|\theta)h(\theta)}{m(x)},$$

gdzie $m(x) = \int_{\Theta} p_\theta(x)h(\theta)d\mu(\theta) = \int_{\Theta} p_\theta(x)d\pi(\theta)$, o ile $m(x) > 0$.

Twierdzenie (Bayes)

Gdy $P(B) > 0$, to $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$.

Twierdzenie (bayesowskie)

Przypuśćmy, że $\pi \in \Pi$ ma gęstość $h(\theta)$ względem miary μ na (Θ, Ξ) .
Dla każdego $x \in X$, istnieje rozkład a posteriori:

$$h(\theta|x) = \frac{p_\theta(x)h(\theta)}{m(x)} = \frac{p(x, \theta|\theta)h(\theta)}{m(x)},$$

gdzie $m(x) = \int_{\Theta} p_\theta(x)h(\theta)d\mu(\theta) = \int_{\Theta} p_\theta(x)d\pi(\theta)$, o ile $m(x) > 0$.

Przykład: $X \sim \mathcal{N}(\theta, \sigma^2)$, $\pi \sim \mathcal{N}(\zeta, \eta^2)$. Wtedy

$$\pi(\cdot|x) \sim \mathcal{N}\left(\frac{x\eta^2 + \zeta\sigma^2}{\sigma^2 + \eta^2}, \frac{\sigma^2\eta^2}{\sigma^2 + \eta^2}\right).$$

Bayesowskie funkcje decyzyjne

Definicja

Ryzykiem oczekiwanym zrandomizowanej funkcji decyzyjnej δ wzgl. rozkł. a priori $\pi \in \Pi$, nazywamy skończoną całkę:

$$r(\pi, \delta) = \mathbb{E}^{\pi} \mathcal{R}(\theta, \delta) = \int_{\Theta} \mathcal{R}(\theta, \delta) d\pi(\theta) .$$

Bayesowskie funkcje decyzyjne

Definicja

Ryzykiem oczekiwanym zrandomizowanej funkcji decyzyjnej δ wzgl. rozkł. a priori $\pi \in \Pi$, nazywamy skończoną całkę:

$$r(\pi, \delta) = \mathbb{E}^{\pi} \mathcal{R}(\theta, \delta) = \int_{\Theta} \mathcal{R}(\theta, \delta) d\pi(\theta) .$$

Funkcja decyzyjna δ^{π} nazywa się bayesowską wzgl. rozkł. a priori π , jeśli $r(\pi) := r(\pi, \delta^{\pi}) = \inf_{\delta \in \mathcal{D}} r(\pi, \delta)$.

Bayesowskie funkcje decyzyjne

Definicja

Ryzykiem oczekiwanym zrandomizowanej funkcji decyzyjnej δ wzgl. rozkł. a priori $\pi \in \Pi$, nazywamy skończoną całkę:

$$r(\pi, \delta) = \mathbb{E}^{\pi} \mathcal{R}(\theta, \delta) = \int_{\Theta} \mathcal{R}(\theta, \delta) d\pi(\theta) .$$

Funkcja decyzyjna δ^{π} nazywa się bayesowską wzgl. rozkł. a priori π , jeśli $r(\pi) := r(\pi, \delta^{\pi}) = \inf_{\delta \in \mathcal{D}} r(\pi, \delta)$.

Jest to równowaga Bayesa — Nasha.

Bayesowskie funkcje decyzyjne

Definicja

Ryzykiem oczekiwanym zrandomizowanej funkcji decyzyjnej δ wzgl. rozkł. a priori $\pi \in \Pi$, nazywamy skończoną całkę:

$$r(\pi, \delta) = \mathbb{E}^{\pi} \mathcal{R}(\theta, \delta) = \int_{\Theta} \mathcal{R}(\theta, \delta) d\pi(\theta) .$$

Funkcja decyzyjna δ^{π} nazywa się bayesowską wzgl. rozkł. a priori π , jeśli $r(\pi) := r(\pi, \delta^{\pi}) = \inf_{\delta \in \mathcal{D}} r(\pi, \delta)$.

Jest to równowaga Bayesa — Nasha.

Twierdzenie

Jeśli istnieje bayesowska funkcja decyzyjna $\delta^{\pi} \in \mathcal{D}$ wzgl. π , to istnieje deterministyczna bayesowska funkcja decyzyjna $d^{\pi} \in \mathcal{D}$.

Bayesowskie funkcje decyzyjne

Definicja

Ryzykiem oczekiwanym zrandomizowanej funkcji decyzyjnej δ wzgl. rozkł. a priori $\pi \in \Pi$, nazywamy skończoną całkę:

$$r(\pi, \delta) = \mathbb{E}^{\pi} \mathcal{R}(\theta, \delta) = \int_{\Theta} \mathcal{R}(\theta, \delta) d\pi(\theta) .$$

Funkcja decyzyjna δ^{π} nazywa się bayesowską wzgl. rozkł. a priori π , jeśli $r(\pi) := r(\pi, \delta^{\pi}) = \inf_{\delta \in \mathcal{D}} r(\pi, \delta)$.

Jest to równowaga Bayesa — Nasha.

Twierdzenie

Jeśli istnieje bayesowska funkcja decyzyjna $\delta^{\pi} \in \mathcal{D}$ wzgl. π , to istnieje deterministyczna bayesowska funkcja decyzyjna $d^{\pi} \in \mathcal{D}$.

Definicja

$\mathcal{R}(\pi(\cdot|x), \delta)$ nazywamy ryzykiem a posteriori.

Gry statystyczne minimaksowe

Definicja

Funkcję decyzyjną $\delta \in \mathcal{D}$ nazywamy minimaxową, jeśli

$$\sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta) = \inf_{d \in \mathcal{D}} \sup_{\theta \in \Theta} \mathcal{R}(\theta, d) =: \bar{v}.$$

Gry statystyczne minimaksowe

Definicja

Funkcję decyzyjną $\delta \in \mathcal{D}$ nazywamy minimaxową, jeśli

$$\sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta) = \inf_{d \in \mathcal{D}} \sup_{\theta \in \Theta} \mathcal{R}(\theta, d) =: \bar{v}.$$

$$(\forall \delta \in \mathcal{D}) \quad \sup_{\pi \in \Pi} r(\pi, \delta) = \sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta)$$

Gry statystyczne minimaksowe

Definicja

Funkcję decyzyjną $\delta \in \mathcal{D}$ nazywamy minimaxową, jeśli

$$\sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta) = \inf_{d \in \mathcal{D}} \sup_{\theta \in \Theta} \mathcal{R}(\theta, d) =: \bar{v}.$$

$$(\forall \delta \in \mathcal{D}) \quad \sup_{\pi \in \Pi} r(\pi, \delta) = \sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta)$$

Wniosek:

Jeśli δ jest minimaksowa w $(\{0, 1\}, \Theta \times \mathcal{D}, -\mathcal{R})$, to też w $(\{0, 1\}, \Theta \times \mathcal{D}, r)$ i $\bar{v}$ jest tożsame.

Gry statystyczne minimaksowe

Definicja

Funkcję decyzyjną $\delta \in \mathcal{D}$ nazywamy minimaxową, jeśli

$$\sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta) = \inf_{d \in \mathcal{D}} \sup_{\theta \in \Theta} \mathcal{R}(\theta, d) =: \bar{v}.$$

$$(\forall \delta \in \mathcal{D}) \quad \sup_{\pi \in \Pi} r(\pi, \delta) = \sup_{\theta \in \Theta} \mathcal{R}(\theta, \delta)$$

Wniosek:

Jeśli δ jest minimaksowa w $(\{0, 1\}, \Theta \times \mathcal{D}, -\mathcal{R})$, to też w $(\{0, 1\}, \Theta \times \mathcal{D}, r)$ i $\bar{v}$ jest tożsame.

Jeśli istnieje wartość gry, to istnieje dokładnie jedna minimaksowa funkcja decyzyjna.