Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie

Nierówności wyrocznie dla problemów odwrotnych

Praca magisterska

Grzegorz Mika

Wydział Matematyki Stosowanej

5 września 2018

Plan prezentacji

- Przedstawienie problemu
- Cel pracy
- Nierówności wyrocznie w modelu z operatorem zwartym
- Nierówności wyrocznie w ogólnym modelu

Przedstawienie problemu

Problem

Niech $A\colon \mathbb{H} \to \mathbb{G}$ będzie operatorem liniowym i ograniczonym między ośrodkowymi przestrzeniami Hilberta. Rozważmy problem

$$Af = g.$$

Cel: mając dany $g \in \mathbb{G}$ znajdź $f \in \mathbb{H}$ by spełnione było powyższe równanie.

Stochastyczny problem odwrotny

Równanie zaburzone przez losowy szum

$$Y = Af + \epsilon\xi, \ \epsilon > 0.$$

Przedstawienie problemu

Dobrze i źle postawione problemy

Problem nazwiemy dobrze postawionym w sensie Hadamarda, gdy:

- dla dowolnego $g \in G$ istnieje $f \in H$ spełniający zadane równanie,
- rozwiązanie jest jedyne,
- rozwiązanie jest stabilne, czyli zależy w sposób ciągły od prawej strony równania.

Problem nazwiemy źle postawionym, gdy nie jest dobrze postawiony.

Błąd stochastyczny

Stochastycznym błędem ξ nazwiemy proces na przestrzeni Hilberta G, czyli ograniczony liniowy operator $\xi\colon G \to L_2(\Omega, \mathcal{F}, \mathbb{P})$ taki, że dla dowolnych elementów $g_1, g_2 \in G$ mamy zdefiniowane zmienne losowe $\langle \xi, g_i \rangle$ takie, że $\mathbb{E}\langle \xi, g_i \rangle = 0$. Operator kowariancji $\mathbf{Cov}_\xi$ określony jest jako ograniczony liniowy operator ($||\mathbf{Cov}_\xi|| \leqslant 1$) z przestrzeni G w przestrzeń G taki, że $\langle \mathbf{Cov}_\xi g_1, g_2 \rangle = \mathbf{Cov}(\langle \xi, g_1 \rangle, \langle \xi, g_2 \rangle)$.

Błąd stochastyczny

Działania na szumie

Niech T będzie operatorem na przestrzeni H. Wtedy można zdefiniować szum $T\xi$ przez

$$\langle T\xi, f\rangle = \langle \xi, T^* f\rangle, \ \forall f \in TH.$$

Operator kowariancji $T\xi$ jest postaci $\mathbf{Cov}_{T\xi} = T\mathbf{Cov}_\xi T^*$.

Estymator liniowy

Estymatorem liniowym elementu f w powyższym modelu nazywamy estymator postaci TY, gdzie T jest pewnym operatorem z przestrzeni $L(G, H)$.

Cel pracy

Niech Λ będzie skończonym zbiorem estymatorów liniowych.

Cel

Wybór z rodziny Λ estymatora naśladującego estymator o minimalnym ryzyku w tej klasie.

Nierówności wyrocznie

Niech $\mathcal{R}(\theta^*, \theta)$ oznacza ryzyko estymatora θ^*. Będziemy poszukiwać metody wyboru θ^* z Λ tak, by udało się skonstruować nierówność następującej postaci

$$\mathcal{R}(\theta^*, \theta) \leqslant C_1 \inf_{\hat{\theta} \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + C_2(\Lambda, n).$$

Model z operatorem zwartym

Niech w problemie $Y = Af + \epsilon\xi$ A będzie operatorem zwartym.

Uwaga

Jeżeli $A\colon \mathbb{H} \to \mathbb{G}$ jest zwarty i $dim\mathbb{H} = \infty$, to A^{-1} jest nieciągły, o ile istnieje, czyli problem $Af = g$ jest źle postawiony dla dowolnego operatora zwartego.

Model z operatorem zwartym

Reprezentacja według wartości singularnych

Istnieją skończony lub zbieżny do zera ciąg liczb dodatnich $\{b_n\}_{n\in I}$ oraz układy ortonormalne $\{v_n\}_{n\in I} \subset H$, $\{u_n\}_{n\in I} \subset G$ takie, że zachodzi

- $KerA^{\perp} = \overline{span\{v_n,\ n \in I\}}$, $\overline{RangeA} = \overline{span\{u_n,\ n \in I\}}$,
- $Af = \sum_n b_n \langle f, v_n \rangle u_n$ oraz $A^* g = \sum_n b_n \langle g, u_n \rangle v_n$.

Przy pewnych dodatkowych warunkach rozwiązania równania $Af = g$ mają postać

$$f = f_0 + \sum_n b_n^{-1} \langle g, u_n \rangle v_n, \ \ f_0 \in KerA.$$

Model z operatorem zwartym

Model ciągowy

Problem $Y = Af + \epsilon\xi$ ma reprezentację postaci

$$x_n = \theta_n + \epsilon\sigma_n\xi_n, \;\; n = 1, 2, \ldots$$

gdzie $x_n = y_n / b_n = \langle Y, u_n \rangle / b_n$, $\theta_n = \langle f, v_n \rangle$ oraz $\sigma_n = b_n^{-1}$.

Estymator liniowy

W powyższym modelu estymator liniowy ma postać $\hat{\theta}(\lambda) = (\hat{\theta}_1, \hat{\theta}_2, \ldots)$, gdzie

$$\hat{\theta}_i = \lambda_i x_i, \;\; i = 1, 2, \ldots$$

dla pewnego nielosowego ciągu liczbowego $\lambda = (\lambda_1, \lambda_2, \ldots)$.

Model z operatorem zwartym

Ryzyko

Ryzyko estymatora θ^* w powyższym modelu wyraża się jako

$$\begin{aligned}\mathcal{R}(\hat{f}, f) &= \mathbb{E}_\theta ||\theta - \hat{\theta}||^2 \\ &= \sum_{n=1}^{\infty} (1-\lambda_n)^2 \theta_n^2 + \epsilon^2 \sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2.\end{aligned}$$

Nieobciążony estymator ryzyka

Nieobciążonym estymatorem $\mathcal{R}(\hat{\theta}, \theta) - \sum_{n=1}^{\infty} \theta_n^2$ nazwiemy wyrażenie

$$U(\lambda, X) = \sum_{n=1}^{\infty} (\lambda_n^2 - 2\lambda_n) x_n^2 + 2\epsilon^2 \sum_{n=1}^{\infty} \lambda_n \sigma_n^2$$

Model z operatorem zwartym

Oznaczenia

$$\rho = \max_{\lambda \in \Lambda} \sup_n \sigma_n^2 |\lambda_n| \left[\sum_{k=1}^{\infty} \sigma_k^4 \lambda_k^4 \right]^{-1/2},$$

$$S = \frac{\max_{\lambda \in \Lambda} \sup_n \sigma_n^2 \lambda_n^2}{\min_{\lambda \in \Lambda} \sup_n \sigma_n^2 \lambda_n^2}, \quad M = \sum_{\lambda \in \Lambda} \exp\left(\frac{-1}{\rho(\lambda)} \right),$$

$$L_\lambda = \ln(NS) + \rho^2 \ln^2(MS).$$

Metoda

$$\lambda^* = \arg\min_{\lambda \in \Lambda} U(\lambda, X)$$

Model z operatorem zwartym

Przy odpowiednich założeniach zachodzą następujące nierówności dla estymatora θ^* wybranego zaproponowaną metodą.

Wyniki

Dla dowolnego $\theta \in l^2$ i dla dowolnego $B > B_0$ istnieją stałe $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ takie, że zachodzi

- $\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant (1 + \gamma_1 B^{-1}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta) + \gamma_2 B \epsilon^2 L_\Lambda \omega(B^2 L_\Lambda)$
- $\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant (1 + \gamma_3 \rho \sqrt{L_\Lambda}) \min_{\lambda \in \Lambda} \mathcal{R}(\hat{\theta}, \theta),$

o ile $\rho\sqrt{L_\Lambda} < \gamma_4$. Funkcja $\omega(x)$ jest postaci

$$\omega(x) = \max_{\lambda \in \Lambda} \sup_k \left[\sigma_k^2 \lambda_k^2 \mathbf{1}\left(\sum_{n=1}^{\infty} \sigma_n^2 \lambda_n^2 \leqslant x \sup_k \sigma_k^2 \lambda_k^2\right)\right], \; x > 0.$$

Model z dowolnym operatorem

Niech w problemie $Y = Af + \epsilon\xi$ A będzie dowolnym operatorem liniowym i ograniczonym.

Prekondycjonowanie problemu

W miejsce problemu

$$Y = Af + \epsilon\xi$$

rozważać będziemy problem

$$A^*Y = A^*Af + \epsilon A^*\xi.$$

Model z dowolnym operatorem

Twierdzenie spektralne

Niech A będzie operatorem samosprzężonym na ośrodkowej przestrzeni Hilberta H. Wtedy istnieją $\sigma-$ zwarta przestrzeń mierzalna $(S, \mathcal{S}, \mu)$ z miarą Radona μ, rzeczywista funkcja mierzalna b określona na S i operator unitarny $U \colon H \to L_2(S, \mathcal{S}, \mu)$, takie, że

$$A = U^{-1} M_b U,$$

gdzie M_b jest operatorem mnożenia przez funkcję b zdefiniowanym jako $(M_b g)(x) = b(x) g(x)$

Model z dowolnym operatorem

Postać

Problem $Y = Af + \epsilon\xi$ można zapisać w postaci

$$X = \theta + \epsilon\sigma\eta,$$

gdzie $\theta = Uf$, $\sigma\eta = M_{b^{-1}}UA^*\xi$ oraz $X = M_{b^{-1}}UA^*Y$.

Estymator

W powyższym modelu estymator liniowy $\hat{\theta}$ przyjmuje postać

$$\hat{\theta} = \lambda X,$$

gdzie λ jest pewną nielosową funkcją.

Model z dowolnym operatorem

Założenia na szum

Załóżmy, że zachodzą następujące warunki gwarantujące skończoność drugich momentów regularyzowanego rozwiązania

$$\forall g \in G \quad \mathbb{E}\langle \xi, g\rangle = 0 \text{ oraz } \|\mathbf{Cov}_\xi\| \leqslant 1,$$

$$\mathbb{E}\,\|A^*\xi\|^2 < \infty.$$

Jeżeli spełnione są powyższe dwa warunki, możliwe jest takie dobranie miary μ i operatora U w twierdzeniu spektralnym, by zachodziło

$$\forall s \in S \quad \mathbf{Var}(UA^*\xi(s)) \leqslant b(s)$$

Model z dowolnym operatorem

Oszacowanie na ryzyko

Przy założeniach na szum ryzyko estymatora $\hat{\theta}$ można oszacować

$$\mathcal{R}(\hat{\theta}, \theta) \leqslant \int_S (1 - \lambda(s))^2 \theta^2(s) d\mu + \epsilon^2 \int_S \lambda^2(s) \sigma(s) d\mu.$$

Oznaczenie

Wprowadźmy oznaczenie

$$\Psi(\lambda, \theta) = \int_S (1 - \lambda(s))^2 \theta^2(s) d\mu + \epsilon^2 \int_S \lambda^2(s) \sigma(s) d\mu.$$

Model z dowolnym operatorem

Empiryczne oszacowanie ryzyka

Wprowadźmy następującą wielkość

$$\psi(\lambda, X) = \int_S (\lambda^2 - 2\lambda) X^2 d\mu + 2\epsilon^2 \int_S \lambda\sigma d\mu.$$

Zachodzi następująca zależność

$$\mathbb{E}\psi(\lambda, X) \leqslant \Psi(\lambda, \theta) - \int_S \theta^2 d\mu.$$

Model z dowolnym operatorem

Metoda

Poszukiwany estymator θ^* będzie konstruowany tak, by filtr był wybierany z rodziny Λ zgodnie z zasadą

$$\lambda^* = \arg\min_{\lambda \in \Lambda} \psi(\lambda, X)$$

Model z dowolnym operatorem

Przy odpowiednich założeniach, oznaczeniach analogicznych do modelu z operatorem zwartym i estymatorze θ^* wybranym zgodnie z zaproponowaną metodą można pokazać następujące nierówności

Wyniki

Dla dowolnego $\theta \in L_2(S, \mathcal{S}, \mu)$, dla dowolnego $B > B_0$ istnieją stałe $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ takie, że zachodzi

- $$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant \max\{1, C_2\}(1 + \gamma_1 B^{-1} \max\{1, \|A\|^2\}) \min_{\lambda \in \Lambda} \Psi(\lambda, \theta) + \max\{1, C_2\}\gamma_2 B \max\{1, \|A\|^2\}\epsilon^2 L_\Lambda \omega(B^2 L_\Lambda)$$
- $$\mathbb{E}_\theta \left\|\theta^* - \theta\right\|^2 \leqslant \max\{1, C_2\}(1 + \gamma_3 \max\{1, \|A\|^2\}\sqrt{\rho L_\Lambda}) \min_{\lambda \in \Lambda} \Psi(\lambda, \theta),$$

o ile $\rho\sqrt{L_\Lambda} < \gamma_4$. Funkcja $\omega(x)$ jest postaci

$$\omega(x) = \max_{\lambda \in \Lambda} \left\|\sigma^2\lambda^2 \mathbf{1}\left(\int_S \sigma\lambda^2 d\mu \leqslant x \left\|\sigma^2\lambda^2\right\|_\infty\right)\right\|_\infty, \; x > 0.$$

Dziękuję za uwagę!