

MARTA LIPNICKA (Łódź)

Przybliżona lokalizacja uszkodzeń w zadanym obszarze

Streszczenie. Niniejsza praca przedstawia procedurę przybliżonego wyznaczania lokalizacji uszkodzenia w zadanym obszarze. Definiujemy problem spektralny, którego rozwiązaniami są wartości własne. Wartości te zależą od położenia i wielkości uszkodzenia. Głównym celem postawionym w pracy jest rozwiązanie zadania odwrotnego, które polega na lokalizowaniu uszkodzenia obszaru na podstawie wektora wartości własnych. Z uwagi na brak jednoznaczności rozwiązań definiujemy nowe zadanie w innym obszarze, dla którego zagwarantowane jest istnienie przybliżonego rozwiązania zadania odwrotnego.

W celu wyznaczenia lokalizacji uszkodzenia definiujemy nowe odwzorowanie, którym jest warunkowa wartość oczekiwana położenia uszkodzenia pod warunkiem, że znamy skończony ciąg wartości własnych. Odwzorowanie to jest aproksymowane przez tak zwany ciąg aproksymujący, którym jest rodzina sieci neuronowych Elmana. Sieć jest budowana w sposób dynamiczny. Jej wielkość jest zależna od liczby elementów należących do zbioru uczącego. Stosowana metoda aproksymacji jest zbieżna.

Słowa kluczowe: Sieci neuronowe, Wartości własne, Aproksymacja, Warunkowa wartość oczekiwana.

1. Wstęp. W otaczającym nas świecie istnieje potrzeba sprawdzania wadliwości różnego rodzaju obiektów. Rozważmy przykład koła kolejowego. Stanowi ono ważny element konstrukcyjny pojazdu szynowego. Co 200 km jazdy pociągu pracownik techniczny kolei ma za zadanie stwierdzić, czy w kołach kolejowych nie powstały jakieś uszkodzenia. Uszkodzeniami mogą być odkształcenia, wygięcia lub inne deformacje widoczne gołym okiem. Oprócz nich w kołach kolejowych mogą powstać szczeliny czy pęknięcia tak małe, że nie są widoczne gołym okiem. W naszych rozważaniach zajmujemy się właśnie takimi deformacjami.

Sprawdzenie koła kolejowego w celu stwierdzenia czy istnieją w nim pęknięcia (szczeliny) polega na tym, że pracownik uderza młotkiem w koło i na podstawie dźwięku jaki otrzymuje jest w stanie określić czy koła mają jakieś uszkodzenia czy też nie. Jeżeli takie uszkodzenie występuje jest bardzo małe, jednak dalsze jego stosowanie grozi powiększeniem defektu i w konsekwencji uszkodzeniem całego koła. Rodzaj dźwięku, który otrzymuje pracownik

przez pukanie młotkiem w koło, jest ściśle zależny od rodzaju defektu, jego wielkości i położenia.

Problemy znajdowania deformacji czy uszkodzenia w materiałach, trybach zębatych, skrzydłach samolotu, itp. stały się bardzo ważne w ostatnich czasach. Procesy produkcyjne stają się coraz bardziej zmechanizowane. Całość produkcji jest realizowana przez taśmy produkcyjne, składające się z różnego rodzaju urządzeń. Rola człowieka ogranicza się w bardzo wielu przypadkach jedynie do nadzorowania poszczególnych etapów produkcji oraz do kontroli prawidłowego jej przebiegu. Kontrola ta polega między innymi na wykrywaniu uszkodzeń w pracujących urządzeniach. Stąd przyrządy do wykrywania defektów poszczególnych elementów tych urządzeń są coraz bardziej potrzebne.

Podobnie jest w innych działach gospodarki. Jako przykład można podać przemysł lotniczy, w którym prawidłowe parametry konstrukcyjne mają kluczowe znaczenie dla eksploatacji samolotów. Jednym z obiektów analiz są parametry skrzydeł samolotów. Koncerny produkujące samoloty kładą duży nacisk na badania naukowe, które dostarczają potrzebnych informacji stosowanych do optymalizacji poszczególnych elementów samolotu oraz do sprawdzania nieprawidłowości w już stosowanym sprzęcie.

W literaturze można znaleźć opisy metod mających na celu wykrywanie deformacji. W pracach [10], [5] opisane są metody szukania deformacji w zadanych obszarach. Deformacjami tymi są między innymi koła. Podobnie w niniejszej pracy szukamy lokalizacji deformacji danego obszaru. W pracy [14] przedstawione jest realizowanie procedury aproksymacji odwzorowania, które jest zdefiniowane jako warunkowa wartość oczekiwana, za pomocą sieci jednokierunkowej wielowarstwowej. Odwzorowanie to może być określone w różny sposób. Może nim być np. przewidywanie pogody na podstawie określonych parametrów takich jak temperatura, ciśnienie atmosferyczne, prędkość wiatru itp. Stosowana przez White'a aproksymacja z wykorzystaniem sieci neuronowej jednokierunkowej (ang. *feed-forward*) jest aproksymacją przez tak zwaną sieć estymatorów. Podobne zastosowania są opisane w pracach [1], [2], [6], [13], [14], [15], [16], [17].

W niniejszej pracy metoda aproksymacji jest realizowana przez sieć estymatorów. Aproksymowane odwzorowanie jest zdefiniowane z wykorzystaniem zadania spektralnego (2). Argumentami odwzorowania jest skończony ciąg wartości własnych wybrany z nieskończonego ciągu rozwiązań zadania spektralnego (2). Wartościami odwzorowania jest warunkowa wartość oczekiwana położenia deformacji, pod warunkiem że znany jest skończony ciąg wartości własnych. Materiał tutaj umieszczony bazuje na [18].

W niniejszej pracy dla wygody rozważamy obszar $\Omega = [0, 1] \times [0, 1]$. Zamiast koła kolejowego obszar poszukiwań stanowi Ω . Szukamy narzędzi, które pozwolą stwierdzić, czy w Ω jest uszkodzenie (szczelina). Badanie de-

formacji – osobliwości na powierzchniach od strony matematycznej przedstawione jest w monografii [9].

W celu stworzenia procedury zmierzającej do wykrycia istnienia deformacji w obszarze Ω definiujemy w tym obszarze problem spektralny, to znaczy rozważamy zadanie

$$(1) \quad \begin{cases} Au_k^0(x_1, x_2) = \lambda_k^0 u_k^0(x_1, x_2) & \text{w } \Omega \\ u_k^0(x_1, x_2) = 0 & \text{na } \partial\Omega \end{cases} \quad k = 1, 2, \dots,$$

gdzie A jest operatorem liniowym $A : H_0^1(\Omega) \mapsto H^{-1}(\Omega)$ takim, że operator $-A$ jest operatorem eliptycznym, a operator odwrotny do A to jest $A^{-1} : L^2(\Omega) \mapsto L^2(\Omega)$ jest operatorem zwartym. Wartości $\lambda_1^0, \lambda_2^0, \dots$ są wartościami własnymi odpowiadającymi obszarowi Ω .

Wartości własne powyższego zadania odnoszą się do obszaru bez deformacji. Jeżeli rozważalibyśmy również podobne zadanie jednak określone dla obszaru z uszkodzeniem, wówczas otrzymalibyśmy, że wartości własne są inne dla obszaru bez dziury i inne dla każdego obszaru ze szczeliną.

Odnosząc się do przypadku koła pociągu, wartości własne rozważanych zadań spektralnych można utożsamiać z dźwiękiem otrzymanym po uderzeniu młotkiem. Jest to inny dźwięk w zależności od rodzaju i położenia deformacji (szczeliny). Stosowane narzędzie, czyli rozwiązanie problemu spektralnego, pozwala na znalezienie relacji między wartościami własnymi (dźwiękiem od uderzenia młotka) a lokalizacją i wielkością szczeliny. Jest to przedmiotem badań niniejszej pracy.

Ponieważ obiektem naszych badań jest szukanie położenia uszkodzenia w zadanym obszarze, w dalszej części pracy będziemy się zajmować problemem ze szczeliną. W obszarze Ω wprowadzamy uszkodzenie (zniekształcenie). Oznaczamy ją przez \mathbf{D} . Ponadto oznaczmy przez $\Omega_{\mathbf{D}}$ obszar Ω z opisaną deformacją. Mamy wówczas $\Omega_{\mathbf{D}} = \Omega \setminus \mathbf{D}$. Dla obszaru $\Omega_{\mathbf{D}}$ rozważamy następujący problem spektralny

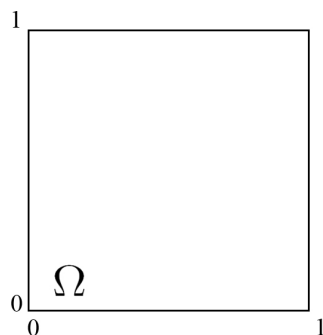
$$(2) \quad \begin{cases} Au_k(x_1, x_2) = \lambda_k u_k(x_1, x_2) & \text{w } \Omega_{\mathbf{D}} \\ u_k(x_1, x_2) = 0 & \text{na } \partial\Omega_{\mathbf{D}} \end{cases} \quad k = 1, 2, \dots$$

Rozwiązaniami powyższego problemu jest nieskończony ciąg par (λ_k, u_k) dla $k = 1, 2, \dots$, gdzie λ_k są wartościami własnymi oraz u_k są funkcjami własnymi. Wartości własne zależą od położenia i wielkości uszkodzenia \mathbf{D} . Dla rozważań prowadzonych w niniejszej pracy istotna jest relacja między wartościami własnymi a wielkością i lokalizacją uszkodzenia.

W rozważanym przez nas przypadku zdefiniowane odwzorowanie posiada niedeterministyczną relację między argumentami a wartościami wyjściowymi. Oznacza to, że nie znamy wartości odwzorowania (wartości opisujące położenie uszkodzenia), która odpowiada ciągowi wartości własnych. Potrafimy jedynie określić warunkową wartość oczekiwaną położenia dziury

pod warunkiem, że znamy ciąg wartości własnych. To powoduje, że musimy użyć narzędzi probabilistycznych do przeprowadzenia procedury aproksymacji.

2. Definicja problemu. Niech dany będzie obszar $\Omega \subset \mathbb{R}^2$, $\Omega = [0, 1] \times [0, 1]$.



Rys. 1. Obszar Ω .

Dla obszaru Ω definiujemy zadanie

$$(3) \quad \begin{cases} Au_k^0(x_1, x_2) = \lambda_k^0 u_k^0(x_1, x_2) & \text{w } \Omega \\ u_k^0(x_1, x_2) = 0 & \text{na } \partial\Omega \end{cases} \quad k = 1, 2, \dots,$$

gdzie A jest operatorem liniowym $A : H_0^1(\Omega) \mapsto H^{-1}(\Omega)$ takim, że operator $-A$ jest operatorem eliptycznym, a operator odwrotny do A tj. $A^{-1} : L^2(\Omega) \mapsto L^2(\Omega)$ jest operatorem zwartym.

Dla powyższego zadania zachodzi twierdzenie.

TWIERDZENIE 1. *Rozważmy zadanie*

$$Au = \lambda u.$$

Ponieważ dla operatora A istnieje operator odwrotny mamy, że

$$u = A^{-1}\lambda u.$$

Z liniowości operatora A^{-1} mamy, że

$$u = \lambda A^{-1}u,$$

a stąd dalej

$$A^{-1}u = \frac{1}{\lambda}u.$$

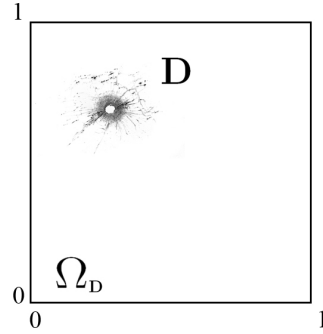
Dla powyższego równania istnieje przeliczalny zbiór wartości własnych

$$\left\{ \frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots \right\}, \quad \lambda_i \neq 0, \quad i \in \mathbb{N},$$

którego jedynym punktem skupienia jest punkt 0. Ponadto każda z wartości własnych jest skończonej krotności.

Rozwiązaniami problemu (3) jest nieskończony ciąg par (λ_k^0, u_k^0) dla $k = 1, 2, \dots$, gdzie λ_k^0 są wartościami własnymi oraz u_k^0 są funkcjami własnymi. Dla wartości własnych mamy, że $0 < \lambda_1^0 < \lambda_2^0 \leq \lambda_3^0 \leq \dots$

W obszarze Ω wprowadzamy uszkodzenie (deformację) \mathbf{D} . Uszkodzeniem jest dziura o nieregularnym kształcie. Może to być szczelina, otwór lub inne zniekształcenie odpowiadające fizycznej deformacji obszaru powstałej na skutek uderzenia, nacisku czy innych sił fizycznych. W praktyce jako obszar Ω możemy przyjąć metalową blaszkę w kształcie kwadratu, w której na skutek uderzenia powstała nieregularna dziura. Zakłada się, że powstała deformacja jest mała (pole powierzchni deformacji jest bliskie 0) i znajduje się wewnątrz zadanego obszaru Ω .



Rys. 2. Obszar $\Omega_{\mathbf{D}}$.

Oznaczmy przez $\Omega_{\mathbf{D}}$ obszar Ω z opisanim uszkodzeniem. Mamy wówczas $\Omega_{\mathbf{D}} = \Omega \setminus \mathbf{D}$. Dla obszaru $\Omega_{\mathbf{D}}$ rozważamy następujący problem spektralny

$$(4) \quad \begin{cases} Au_k(x_1, x_2) = \lambda_k u_k(x_1, x_2) & \text{w } \Omega_{\mathbf{D}} \\ u_k(x_1, x_2) = 0 & \text{na } \partial\Omega_{\mathbf{D}} \end{cases} \quad k = 1, 2, \dots$$

Rozwiązaniami powyższego problemu jest nieskończony ciąg par (λ_k, u_k) dla $k = 1, 2, \dots$, gdzie λ_k są wartościami własnymi oraz u_k są funkcjami własnymi. Wartości własne zależą od położenia i wielkości zniekształcenia.

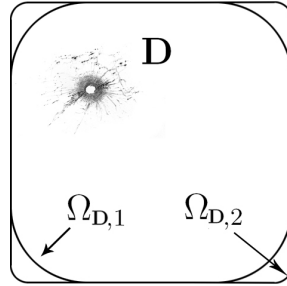
Wartości własne są opisem pewnych własności fizycznych obszaru. Mogą nimi być np. temperatura czy dźwięk. Chcemy znaleźć zależność między nieskończonym ciągiem wartości własnych a zniekształceniem obszaru. Innymi słowy mając dany ciąg wartości własnych chcemy lokalizować uszkodzenie przez znalezienie dwuwymiarowego wektora będącego wartością oczekiwaną położenia zniekształcenia. W omawianym zadaniu nie mamy jednoznaczności rozwiązań. Oznacza to, że nie można w sposób jednoznaczny na podstawie wartości własnych wyznaczyć położenia zniekształcenia. Istnieją bowiem

różne deformacje obszaru, dla których wartości własne są identyczne.

Wprowadzamy aproksymację obszaru $\Omega_{\mathbf{D}}$ ciągiem obszarów $\Omega_{\mathbf{D},n}$, gdzie $n = 1, 2, \dots$. Ciąg obszarów $\Omega_{\mathbf{D},n}$ jest zdefiniowany w taki sposób, że dla każdego $n = 1, 2, \dots$ deformacja \mathbf{D} jest taka sama (nie ulega zmianie) oraz dla każdego $n = 1, 2, \dots$ $\Omega_{\mathbf{D},n} \subset \Omega_{\mathbf{D}}$ i $\partial\Omega_{\mathbf{D},n} \rightarrow \partial\Omega_{\mathbf{D}}$ gdy $n \rightarrow \infty$. Innymi słowy ciąg obszarów aproksymujących $\Omega_{\mathbf{D}}$ zdefiniowany jest w taki sposób, że dla każdego $n = 1, 2, \dots$ $\partial\Omega_{\mathbf{D},n} = G_n \cup \partial\mathbf{D}$, gdzie \mathbf{D} jest taką samą deformacją jak dla $\Omega_{\mathbf{D}}$ oraz

$$(5) \quad G_n$$

jest zbiorem wartości pewnej krzywej zamkniętej g_n . Zbiory wartości krzywych zamkniętych g_n , $n = 1, 2, \dots$ są zdefiniowane w taki sposób, że $G_n \rightarrow \partial([0, 1] \times [0, 1])$ przy $n \rightarrow \infty$. Ponadto g_n jest klasy C^2 .



Rys. 3. Ciąg obszarów $\Omega_{\mathbf{D},n}$, $n = 1, 2, \dots$

Wprowadzona powyżej konstrukcja brzegu $\Omega_{\mathbf{D},n}$ oznacza, że dla każdych $n, m = 1, 2, \dots$ takich, że $n < m$, $\Omega_{\mathbf{D},n} \subset \Omega_{\mathbf{D},m}$. Należy podkreślić, że w każdym obszarze $\Omega_{\mathbf{D},n}$ zniekształcenie \mathbf{D} jest takie samo. Ma to kluczowe znaczenie przy rozważaniu problemu odwrotnego, czyli problemu, w którym dla zadanego obszaru szukamy wartości oczekiwanej położenia deformacji. Z uwagi na to, że dla $n = 1, 2, \dots$ każdy z obszarów $\Omega_{\mathbf{D},n}$ ma takie samo zniekształcenie \mathbf{D} , rozwiązanie zadania odwrotnego dla dowolnego n zwraca ten sam rezultat. Innymi słowy, jeżeli dla dowolnego (ale ustalonego) n na podstawie wartości własnych wyliczymy wartość oczekiwaną położenia zniekształcenia \mathbf{D} , to otrzymany rezultat będzie jednocześnie rozwiązaniem dla pozostałych obszarów $\Omega_{\mathbf{D},n}$ oraz rozwiązaniem problemu spektralnego (4) dla obszaru wyjściowego $\Omega_{\mathbf{D}}$.

3. Zadanie odwrotne. Niech dany będzie obszar $\Omega \subset \mathbb{R}^2$, $\Omega = [0, 1] \times [0, 1]$ i rozważane wcześniej zadanie wyjściowe

$$\begin{cases} Au_k^0(x_1, x_2) = \lambda_k^0 u_k^0(x_1, x_2) & \text{w } \Omega \\ u_k^0(x_1, x_2) = 0 & \text{na } \partial\Omega \end{cases} \quad k = 1, 2, \dots$$

Rozwiązaniem powyższego problemu jest nieskończony ciąg par (λ_k^0, u_k^0) dla $k = 1, 2, \dots$, gdzie λ_k^0 są wartościami własnymi oraz u_k^0 są funkcjami własnymi.

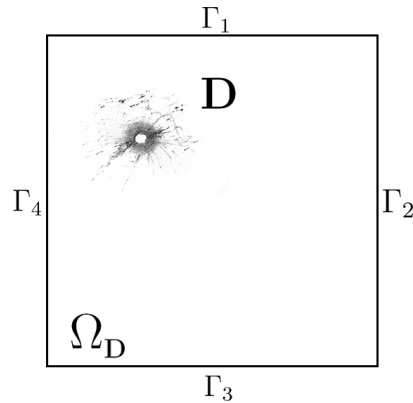
Niech dany będzie ponadto obszar $\Omega_D = \Omega \setminus D$ i rozważane wcześniej zadanie

$$\begin{cases} Au_k(x_1, x_2) = \lambda_k u_k(x_1, x_2) & \text{w } \Omega_D \\ u_k(x_1, x_2) = 0 & \text{na } \partial\Omega_D \end{cases} \quad k = 1, 2, \dots$$

Rozwiązaniami powyższego problemu jest nieskończony ciąg par (λ_k, u_k) dla $k = 1, 2, \dots$, gdzie λ_k są wartościami własnymi oraz u_k są funkcjami własnymi. Wartości rozwiązań zadania (4) zależą od wielkości i położenia zniekształcenia D . Weźmy pod uwagę jedynie wartości własne i ustalmy zależność między $\lambda_k, k = 1, 2, \dots$ a D . Ponieważ wartości $\lambda_k, k = 1, 2, \dots$ zależą od wielkości i położenia zniekształcenia D możemy określić funkcję, która każdej deformacji przyporządkowuje nieskończony ciąg wartości własnych: $D \rightarrow (\lambda_1, \lambda_2, \dots)$. Tak zdefiniowana funkcja nie jest różnowartościowa. Istnieje kilka różnych deformacji, dla których ciąg wartości własnych $(\lambda_1, \lambda_2, \dots)$ jest taki sam.

Jako główny cel stawiamy sobie rozwiązanie zadania odwrotnego, które polega na lokalizowaniu zniekształcenia obszaru Ω_D na podstawie wektora wartości własnych (podobne zastosowanie patrz [8]). Z uwagi na brak jednoznaczności rozwiązań zadania (4) musimy zdefiniować nowe zadanie w innym obszarze, który zagwarantuje nam istnienie przybliżonego rozwiązania zadania odwrotnego.

Niech dany będzie obszar Ω_D . Wprowadźmy następujące oznaczenia: $\Gamma_1 = [0, 1] \times \{1\}$, $\Gamma_2 = \{1\} \times [0, 1]$, $\Gamma_3 = [0, 1] \times \{0\}$ oraz $\Gamma_4 = \{0\} \times [0, 1]$.

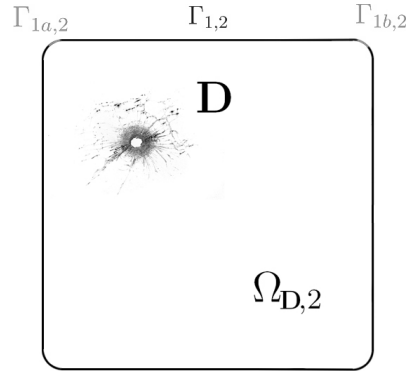


Rys. 4. Obszar Ω_D .

Obszar Ω_D jest aproksymowany przez ciąg obszarów $\Omega_{D,n}, n = 1, 2, \dots$. Niech $\Gamma_{1,n}, \Gamma_{2,n}, \Gamma_{3,n}, \Gamma_{4,n}, n = 1, 2, \dots$ będą fragmentami $\partial\Omega_D$ wspólnymi

odpowiednio z $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$. Oznacza to, że $\Gamma_{1,n} = \Gamma_1 \cap G_n, \dots, \Gamma_{4,n} = \Gamma_4 \cap G_n$, gdzie G_n jest zbiorem wartości pewnej krzywej zamkniętej g_n (patrz oznaczenie (5)).

Niech ponadto $\Gamma_{1a,n}, \Gamma_{1b,n}, \Gamma_{2a,n}, \Gamma_{2b,n}, \Gamma_{3a,n}, \Gamma_{3b,n}, \Gamma_{4a,n}, \Gamma_{4b,n}$, $n = 1, 2, \dots$ oznaczają te fragmenty $\partial\Omega_{\mathcal{D}}$, które zostały zmodyfikowane odpowiednio w $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$. Oznacza to, że $\Gamma_{1a,n}$ jest zbiorem wartości krzywej, która powstaje ze zbioru wartości funkcji h_n , obciętej do przedziału $[0, \frac{1}{2^{n-1}\sqrt{[4]2\pi e}}]$, po obrocie o kąt $\frac{7\pi}{4}$. Bardziej formalnie możemy zapisać, że $\Gamma_{1a,n} = h_n |_{[0, \frac{1}{2^{n-1}\sqrt{[4]2\pi e}}]}$ ⁽¹⁾ po obrocie o kąt $\frac{7\pi}{4}$. Analogicznie $\Gamma_{1b,n} = h_n |_{[-\frac{1}{2^{n-1}\sqrt{[4]2\pi e}}, 0]}$ po obrocie o kąt $\frac{\pi}{4}$. W ten sam sposób można zdefiniować $\Gamma_{2a,n}, \Gamma_{2b,n}, \Gamma_{3a,n}, \Gamma_{3b,n}, \Gamma_{4a,n}, \Gamma_{4b,n}$, $n = 1, 2, \dots$



Rys. 5. Obszar $\Omega_{\mathcal{D},2}$.

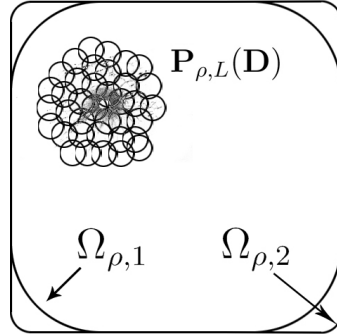
Konstrukcja $\Omega_{\mathcal{D},n}$, $n = 1, 2, \dots$ zakłada, że zniekształcenie każdego z obszarów jest takie samo jak zniekształcenie obszaru $\Omega_{\mathcal{D}}$ (to znaczy \mathbf{D}). Niech ρ będzie ustaloną liczbą rzeczywistą dodatnią, bliską 0. Ustalmy liczbę $L \in \mathbb{N}$. Liczba L jest dobrana niezależnie od rodzaju zniekształcenia. Jest na tyle duża, że każde zniekształcenie może być pokryte L kołami, każde o promieniu ρ . Ustalenie liczby L jest możliwe ze względu na ograniczoność obszaru Ω i założenie, że zniekształcenie obszaru jest małe.

Niech dane będzie pokrycie zniekształcenia zdefiniowane jako skończony ciąg kół $B_l(\bar{y}_l, \rho) = \{y \in (0, 1) \times (0, 1) : |y - \bar{y}_l| < \rho\}$, gdzie $l = 1, 2, \dots, L$. Zauważmy, że $\partial\{\bigcup_{l=1}^L B_l(\bar{y}_l, \rho)\}$ może być niegładki. Oznaczmy przez $\mathbf{P}_{\rho,L}(\mathbf{D})$ zmodyfikowany zbiór $\bigcup_{l=1}^L B_l(\bar{y}_l, \rho)$ w taki sposób, że $\partial\mathbf{P}_{\rho,L}(\mathbf{D})$ jest gładki oraz $\mathbf{P}_{\rho,L}(\mathbf{D})$ jest pokryciem zbioru \mathbf{D} zawierającym $\bigcup_{l=1}^L B_l(\bar{y}_l, \rho)$ i dowolnie bliskim $\bigcup_{l=1}^L B_l(\bar{y}_l, \rho)$. W przypadku, gdy deformacja \mathbf{D} jest bardzo

⁽¹⁾ Zapis $h_n |_{[0, \frac{1}{2^{n-1}\sqrt{[4]2\pi e}}]}$ oznacza obcięcie funkcji h_n do przedziału $[0, \frac{1}{2^{n-1}\sqrt{[4]2\pi e}}]$

mała możemy przyjąć, że rozważanym dla niej pokryciem będzie jedna kula. Wówczas $\mathbf{P}_{\rho,L}(\mathbf{D}) = B_1(\bar{y}_1, \rho)$.

Oznaczmy przez Ω_ρ obszar Ω z deformacją, którą w tym przypadku jest $\mathbf{P}_{\rho,L}(\mathbf{D})$. Mamy wówczas $\Omega_\rho = \Omega \setminus \mathbf{P}_{\rho,L}(\mathbf{D})$. Oznaczmy analogicznie $\Omega_{\rho,n} = \Omega_{\mathbf{D},n} \setminus \mathbf{P}_{\rho,L}(\mathbf{D})$. $\Omega_{\rho,n}$ jest zdefiniowane w taki sposób, że $\partial\Omega_{\rho,n}$ jest gładki. Ponadto $\Omega_{\rho,n} \rightarrow \Omega_{\mathbf{D}}$, przy $n \rightarrow \infty$ i $\rho \rightarrow 0$.



Rys. 6. Obszary $\Omega_{\rho,n}$, $n = 1, 2$.

Dla ustalonych n i L definiujemy zadanie

$$(6) \quad \begin{cases} A\bar{u}_{k,n}(x_1, x_2) = \lambda_k \bar{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \bar{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{1a,n}, \quad k = 1, 2, \dots \\ \bar{u}_{k,n}(x_1, x_2) = u_k(x_1, x_2) & \text{na } \Gamma_{1a,n}. \end{cases}$$

gdzie u_k , λ_k są rozwiązaniami (4). Zauważmy, że warunek brzegowy nie jest funkcją ciągłą. Rozważmy kolejny problem

$$(7) \quad \begin{cases} A\bar{u}_{k,n}(x_1, x_2) = \lambda_k \bar{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \bar{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{1a,n}, \quad k = 1, 2, \dots \\ \bar{u}_{k,n}(x_1, x_2) = \alpha_k(x_1, x_2) & \text{na } \Gamma_{1a,n}. \end{cases}$$

gdzie λ_k są rozwiązaniami (4) zaś α_k są funkcjami klasy C^2 dowolnie bliskimi u_k , gdzie u_k są rozwiązaniami (4). Ponadto α_k są zdefiniowane w taki sposób, że warunek brzegowy (7) jest funkcją ciągłą.

Rozwiązaniami (7) są funkcje $\bar{u}_{k,n}$, $k = 1, 2, \dots$. Wartości λ_k , $k = 1, 2, \dots$ są rozwiązaniami (4). Zarówno λ_k , jak i $\bar{u}_{k,n}$, $k = 1, 2, \dots$ zależą od wielkości i położenia zniekształcenia \mathbf{D} , które od początku naszych rozważań nie zmieniło się.

Niech X i Y będą wektorami zdefiniowanymi następująco.

$$(8) \quad X = [\lambda_1, \dots, \lambda_r],$$

gdzie $r \in \mathbb{N}$ określa ustaloną liczbę wartości własnych. X jest wektorem r -elementowym, którego składowymi jest r wartości własnych będących rozwiązaniami (4). Wartości te są ustalone, to znaczy na początku definicji wek-

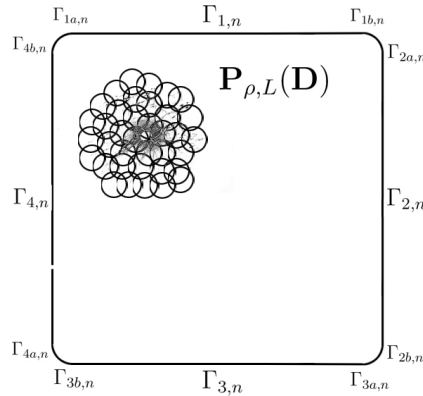
tora X ustala się indeksy (niekoniecznie kolejne) wartości własnych, które będą go definiować. Y jest wektorem określającym położenie zniekształcenia D w obszarze $\Omega_{\rho,n}$. Y jest wyznaczany następująco. Pokryciem zniekształcenia D jest $P_{\rho,L}(D)$. $P_{\rho,L}(D)$ składa się z L kół $B_l(\bar{y}_l, \rho)$, każde o promieniu ρ i środku \bar{y}_l . Każdemu z kół przypisujemy liczby $p_l^1, p_l^2, l = 1, 2, \dots, L$ takie, że dla każdego $l = 1, 2, \dots, L, p_l^1 > 0, p_l^2 > 0$ oraz $\sum_{l=1}^L p_l^1 = 1, \sum_{l=1}^L p_l^2 = 1$. Ponadto wartości p_l^1 i p_l^2 są dobrane tak, że im większa ich część pokrywa się ze zniekształceniem tym wyższe są wartości p_l^1 oraz p_l^2 . Pokrycie $P_{\rho,L}(D)$ jest zdefiniowane w taki sposób, że wszystkie zawarte w nim koła mają ten sam promień. Mamy, że zniekształcenie D składa się z fragmentów różnej szerokości. Zachodzi zatem zależność $D \subset \bigcup_{l=1}^L B_l(\bar{y}_l, \rho) \subset P_{\rho,L}(D)$ oraz $\bigcup_{l=1}^L B_l(\bar{y}_l, \rho) \cap \Omega_{D,n} \neq \emptyset$. Tym samym będą takie koła z pokrycia zniekształcenia D , których bardzo mała część pola pokrywa zniekształcenie, natomiast pozostała część pokrywa fragment $\Omega_{D,n}$ przyległy zniekształceniu D . Dla takich kół liczby p_l^1, p_l^2 są bliskie 0. Y jest wyliczane na podstawie pokrycia w następujący sposób.

$$(9) \quad Y = [Y^1, Y^2],$$

gdzie $Y^1 = \sum_{l=1}^L (p_l^1 \cdot \bar{y}_l^1), Y^2 = \sum_{l=1}^L (p_l^2 \cdot \bar{y}_l^2)$.

3.1. Jednoznaczność rozwiązań. Niech dany będzie obszar $\Omega \subset \mathbb{R}^2, \Omega = [0, 1] \times [0, 1]$. Przypomnijmy, że $\Gamma_1 = [0, 1] \times \{1\}, \Gamma_2 = \{1\} \times [0, 1], \Gamma_3 = [0, 1] \times \{0\}$ oraz $\Gamma_4 = \{0\} \times [0, 1]$.

Rozważmy wprowadzone wcześniej D oraz odpowiadający mu obszar Ω_ρ . Stosujemy aproksymację $\partial\Omega_\rho$ i otrzymujemy ciąg obszarów $\Omega_{\rho,n}, n = 1, 2, \dots$



Rys. 7. Obszar $\Omega_{\rho,n}$, dla ustalonego n .

Dla ustalonych n i L definiujemy dwa nowe zadania. Pierwsze zadanie

odpowiada (6) i jest zdefiniowane następująco

$$(10) \quad \begin{cases} -A\check{u}_{k,n}(x_1, x_2) = \lambda_k \check{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \check{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{1a,n}, \\ \check{u}_{k,n}(x_1, x_2) = u_k(x_1, x_2) & \text{na } \Gamma_{1a,n}. \end{cases} \quad k = 1, 2, \dots$$

gdzie u_k , λ_k są rozwiązaniami (4); oraz drugie – odpowiadające (7) zdefiniowane następująco

$$(11) \quad \begin{cases} -A\check{u}_{k,n}(x_1, x_2) = \lambda_k \check{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \check{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{1a,n}, \\ \check{u}_{k,n}(x_1, x_2) = \alpha_k(x_1, x_2) & \text{na } \Gamma_{1a,n}, \end{cases} \quad k = 1, 2, \dots$$

gdzie λ_k , tak jak poprzednio, są rozwiązaniami (4), zaś α_k są funkcjami klasy C^2 dowolnie bliskimi u_k , gdzie u_k są rozwiązaniami (4). Ponadto α_k są zdefiniowane w taki sposób, że warunek brzegowy jest funkcją ciągłą. Zadanie (11) jest zadaniem pomocniczym.

Dla ustalonych $k \in \{1, \dots, r\}$ rozważmy następujący problem

$$(12) \quad \begin{cases} -A\tilde{u}_{k,n}(x_1, x_2) = \lambda_k \tilde{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \tilde{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{1a,n}, \\ \tilde{u}_{k,n}(x_1, x_2) = \beta_k(x_1, x_2) & \text{na } \Gamma_{1a,n}, \end{cases} \quad k = 1, \dots, r,$$

gdzie k oznacza indeks ustalonych wartości własnych, które zostały wybrane do zdefiniowania wektora X , zaś funkcje β_k są funkcjami nieujemnymi, zdefiniowanymi w taki sposób, że warunek brzegowy jest funkcją ciągłą oraz spełnia $|\beta_k| < \frac{1}{n}$ dla $k = 1, \dots, r$.

Rozważmy ponadto taki sam problem, który powstaje przez obrót obszaru $\Omega_{\rho,n}$ o kąt $\frac{\pi}{2}$ względem punktu $(\frac{1}{2}, \frac{1}{2})$.

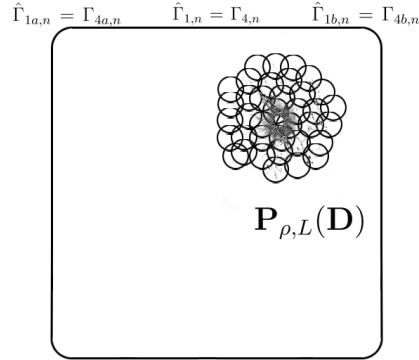
$$(13) \quad \begin{cases} -A\hat{u}_{k,n}(x_1, x_2) = \lambda_k \hat{u}_{k,n}(x_1, x_2) & \text{w } \Omega_{\rho,n}, \\ \hat{u}_{k,n}(x_1, x_2) = 0 & \text{na } \partial\Omega_{\rho,n} \setminus \Gamma_{4a,n}, \\ \hat{u}_{k,n}(x_1, x_2) = \gamma_k(x_1, x_2) & \text{na } \Gamma_{4a,n}, \end{cases} \quad k = 1, \dots, r,$$

gdzie k oznacza indeks ustalonych wartości własnych, które zostały wybrane do zdefiniowania wektora X , zaś funkcje γ_k są funkcjami nieujemnymi, zdefiniowanymi w taki sposób, że warunek brzegowy jest funkcją ciągłą oraz spełnia $|\gamma_k| < \frac{1}{n}$ dla $k = 1, \dots, r$.

Możemy wprowadzić notację

$$\begin{aligned} \hat{\Gamma}_{1,n} &= \Gamma_{4,n}, \hat{\Gamma}_{2,n} = \Gamma_{1,n}, \hat{\Gamma}_{3,n} = \Gamma_{2,n}, \hat{\Gamma}_{4,n} = \Gamma_{3,n}, \\ \hat{\Gamma}_{1a,n} &= \Gamma_{4a,n}, \hat{\Gamma}_{2a,n} = \Gamma_{1a,n}, \hat{\Gamma}_{3a,n} = \Gamma_{2a,n}, \hat{\Gamma}_{4a,n} = \Gamma_{3a,n}, \\ \hat{\Gamma}_{1b,n} &= \Gamma_{4b,n}, \hat{\Gamma}_{2b,n} = \Gamma_{1b,n}, \hat{\Gamma}_{3b,n} = \Gamma_{2b,n}, \hat{\Gamma}_{4b,n} = \Gamma_{3b,n}. \end{aligned}$$

Załóżmy, że mamy dane λ , które jest ustaloną wartością własną, będącą rozwiązaniem (4) i stanowi składową wektora X . Dla ustalonego $\Omega_{\rho,n}$ na podstawie Twierdzenia 6.25 z [4] wiemy, że istnieje dokładnie jedno rozwiązanie \tilde{u} i \hat{u} odpowiednio problemu (12) i (13). Zadanie jest rozważane



Rys. 8. Obszar $\Omega_{\rho,n}$, dla ustalonego n po obrocie o kąt $\frac{\pi}{2}$ względem punktu $(\frac{1}{2}, \frac{1}{2})$.

dla każdego k osobno. Ponieważ n i k zostały ustalone wcześniej w celu uproszczenia zapisu indeksy przy funkcjach są pominięte.

Niech $y = (y_1, y_2)$ będzie dwuelementowym wektorem, który opisuje położenie zniekształcenia \mathbf{D} . Przyjmujemy, że $y = Y$, gdzie Y dany jest wzorem (9). Mając dane \tilde{u} i \hat{u} możemy udowodnić następujące twierdzenie.

TWIERDZENIE 2. *Niech \tilde{u} i \hat{u} będą odpowiednio rozwiązaniami (12) i (13). Wówczas możemy w sposób jednoznaczny wyznaczyć wektor $y = (y_1, y_2)$, który określa położenie zniekształcenia \mathbf{D} w obszarze $\Omega_{\rho,n}$.*

Dowód. Rozważmy całkę

$$(14) \quad \mathbb{I} = \int_{\Omega_{\rho,n}} \operatorname{div}(\tilde{u}(x) x) d\Omega_{\rho,n},$$

gdzie $x = [x_1, x_2]$.

Nie zmniejszając ogólności rozważań przesuujemy początek układu współrzędnych Ox_1x_2 do punktu $y = (y_1, y_2)$. Wówczas parametryzacja $\Gamma_{1,n}, \Gamma_{2,n}, \Gamma_{3,n}, \Gamma_{4,n}$ ma następującą postać

$$x_2 = 1 - y_2 \text{ na } \Gamma_{1,n},$$

$$x_1 = 1 - y_1 \text{ na } \Gamma_{2,n},$$

$$x_2 = -y_2 \text{ na } \Gamma_{3,n},$$

$$x_1 = -y_1 \text{ na } \Gamma_{4,n}.$$

Powyższy zapis jest poprawny, gdyż dla

$$\Gamma_{1,n} \subset \Gamma_1 = [0, 1] \times \{1\} \text{ i } \Gamma_{3,n} \subset \Gamma_3 = [0, 1] \times \{0\}$$

druga współrzędna jest stała, zaś dla

$$\Gamma_{2,n} \subset \Gamma_2 = \{1\} \times [0, 1] \text{ oraz dla } \Gamma_{4,n} \subset \Gamma_4 = \{0\} \times [0, 1]$$

pierwsza współrzędna jest stała.

Przyjmujemy, że zapis \tilde{u} oznacza funkcję \tilde{u} zmiennej x . Stosując twier-

dzenie o dywergencji dla wyrażenia (14) otrzymujemy, że

$$\mathbb{I} = \int_{\partial\Omega_{\rho,n}} \tilde{u}(x) (x \cdot \mathbf{n}) dS,$$

gdzie $\mathbf{n} = (n_1, n_2)$ jest wektorem normalnym do brzegu $\partial\Omega_{\rho,n}$ i ma długość 1 zaś operator $'\cdot'$ oznacza iloczyn skalarny wektorów. Z własności geometrycznych poszczególnych fragmentów brzegu $\partial\Omega_{\rho,n}$ mamy, że

$$x \cdot \mathbf{n} = x_2 = 1 - y_2 \text{ na } \Gamma_{1,n},$$

$$x \cdot \mathbf{n} = x_1 = 1 - y_1 \text{ na } \Gamma_{2,n},$$

$$x \cdot \mathbf{n} = x_2 = -y_2 \text{ na } \Gamma_{3,n},$$

$$x \cdot \mathbf{n} = x_1 = -y_1 \text{ na } \Gamma_{4,n}.$$

Wówczas

$$\begin{aligned} \mathbb{I} &= (1 - y_2) \int_{\Gamma_{1,n}} \tilde{u} dS + (1 - y_2) \left(\int_{\Gamma_{1a,n}} \tilde{u} dS + \int_{\Gamma_{1b,n}} \tilde{u} dS \right) \\ &\quad + \left(\int_{\Gamma_{1a,n}} w_{1,n} \tilde{u} dS + \int_{\Gamma_{1b,n}} w_{1,n} \tilde{u} dS \right) \\ &\quad + (1 - y_1) \int_{\Gamma_{2,n}} \tilde{u} dS + (1 - y_1) \left(\int_{\Gamma_{2a,n}} \tilde{u} dS + \int_{\Gamma_{2b,n}} \tilde{u} dS \right) \\ &\quad + \left(\int_{\Gamma_{2a,n}} w_{2,n} \tilde{u} dS + \int_{\Gamma_{2b,n}} w_{2,n} \tilde{u} dS \right) \\ &\quad + y_2 \int_{\Gamma_{3,n}} \tilde{u} dS + y_2 \left(\int_{\Gamma_{3a,n}} \tilde{u} dS + \int_{\Gamma_{3b,n}} \tilde{u} dS \right) \\ &\quad + \left(\int_{\Gamma_{3a,n}} w_{3,n} \tilde{u} dS + \int_{\Gamma_{3b,n}} w_{3,n} \tilde{u} dS \right) \\ &\quad + y_1 \int_{\Gamma_{4,n}} \tilde{u} dS + y_1 \left(\int_{\Gamma_{4a,n}} \tilde{u} dS + \int_{\Gamma_{4b,n}} \tilde{u} dS \right) \\ &\quad + \left(\int_{\Gamma_{4a,n}} w_{4,n} \tilde{u} dS + \int_{\Gamma_{4b,n}} w_{4,n} \tilde{u} dS \right) + \int_{\partial\mathcal{P}_{\rho,L}(\mathcal{D})} \tilde{u}(x) (x \cdot \mathbf{n}) dS, \end{aligned}$$

gdzie $w_{1,n}$, $w_{2,n}$, $w_{3,n}$ i $w_{4,n}$ są funkcjami ciągłymi takimi, że

$$(1 - y_2) + w_{1,n}(x) = x \cdot \mathbf{n} \text{ po } \Gamma_{1a,n} \cup \Gamma_{1b,n},$$

$$(1 - y_1) + w_{2,n}(x) = x \cdot \mathbf{n} \text{ po } \Gamma_{2a,n} \cup \Gamma_{2b,n},$$

$$y_2 + w_{3,n}(x) = x \cdot \mathbf{n} \text{ po } \Gamma_{3a,n} \cup \Gamma_{3b,n} \text{ i}$$

$$y_1 + w_{4,n}(x) = x \cdot \mathbf{n} \text{ po } \Gamma_{4a,n} \cup \Gamma_{4b,n}.$$

Funkcje $w_{1,n}$, $w_{2,n}$, $w_{3,n}$ i $w_{4,n}$ są definiowane dla każdego n osobno, gdyż ich konstrukcja zależy od brzegu obszaru.

Z własności funkcji \tilde{u} na $\partial\Omega_{\rho,n} \setminus \Gamma_{1,n}$ mamy, że

$$\mathbb{I} = (1 - y_2) \left(\int_{\Gamma_{1a,n}} \tilde{u} \, dS \right) + \left(\int_{\Gamma_{1a,n}} w_{1,n} \tilde{u} \, dS \right).$$

Przekształcając powyższą równość mamy dalej

$$-y_2 \left(\int_{\Gamma_{1a,n}} \tilde{u} \, dS \right) = \mathbb{I} - \left(\int_{\Gamma_{1a,n}} \tilde{u} \, dS \right) - \left(\int_{\Gamma_{1a,n}} w_{1,n} \tilde{u} \, dS \right).$$

Stosując analogiczne rozumowanie do zadania (13) otrzymujemy następujące równanie

$$(15) \quad -\hat{y}_2 \left(\int_{\hat{\Gamma}_{1a,n}} \hat{u} \, dS \right) = \mathbb{I} - \left(\int_{\hat{\Gamma}_{1a,n}} \hat{u} \, dS \right) - \left(\int_{\hat{\Gamma}_{1a,n}} \hat{w}_{1,n} \hat{u} \, dS \right).$$

Ponieważ $\hat{\Gamma}_{1a,n} = \Gamma_{4a,n}$, $\hat{w}_{1,n} = w_{4,n}$ oraz $\hat{y}_2 = 1 - y_1$ otrzymujemy dalej

$$(16) \quad \begin{aligned} -(1 - y_1) \left(\int_{\Gamma_{4a,n}} \hat{u} \, dS \right) &= \mathbb{I} - \left(\int_{\Gamma_{4a,n}} \hat{u} \, dS \right) - \left(\int_{\Gamma_{4a,n}} w_{4,n} \hat{u} \, dS \right) \\ y_1 \left(\int_{\Gamma_{4a,n}} \hat{u} \, dS \right) &= \mathbb{I} - \left(\int_{\Gamma_{4a,n}} w_{4,n} \hat{u} \, dS \right). \end{aligned}$$

Traktujemy (15) oraz (16) jako układ dwóch równań z dwiema niewiadomymi y_1 i y_2 . Wyznacznik W tego układu przyjmuje postać.

$$W = \int_{\Gamma_{4a,n}} \hat{u} \, dS \int_{\Gamma_{1a,n}} \tilde{u} \, dS.$$

Z własności funkcji \hat{u} i \tilde{u} odpowiednio na $\Gamma_{4a,n}$ i $\Gamma_{1a,n}$ mamy, że $W \neq 0$, a to daje nam tezę twierdzenia. \square

Powyższe postępowanie pokazuje, jak na podstawie ustalonej wartości własnej wyznaczyć w sposób jednoznaczny położenie zniekształcenia \mathbf{D} w ustalonym obszarze $\Omega_{\rho,n}$. W naszym przypadku nie będzie to dokładne, a jedynie przybliżone położenie zniekształcenia. Jest to związane z wyznaczeniem funkcji \tilde{u} (analogicznie dla \hat{u}). Procedura wyznaczenia położenia uszkodzenia jest przeprowadzana w dwóch krokach. Pierwszy polega na wyznaczeniu na podstawie danej wartości własnej, będącej rozwiązaniem (4), funkcji własnej \tilde{u} . Funkcja ta jest rozwiązaniem problemu (12). Jej istnienie gwarantuje twierdzenie 6.25 z [4] cytowane wcześniej. Dokładne rozwiązanie problemu (12) jest bardzo trudne, a w niektórych przypadkach niemożliwe. Stąd wartości wejściowe do powyższego twierdzenia są wartościami przybliżonymi i na ich podstawie wyznaczane jest (również przybliżone) położenie zniekształcenia. Ponadto funkcja \tilde{u} nie jest funkcją własną odpowiadającą

wartości własnej zadania (4) ani problemu (12). Jest tylko jej pewnym, prawdopodobnym przybliżeniem. Drugi krok, to wyznaczenie na podstawie funkcji \tilde{u} wektora określającego położenie zniekształcenia \mathbf{D} . Ponieważ funkcja \tilde{u} jest podana z pewnym przybliżeniem, to otrzymana lokalizacja zniekształcenia również jest przybliżona.

Ponieważ wartość własna służąca do definicji zadania (12) jest rozwiązaniem zadania (4) można zauważyć, że otrzymany rezultat lokalizacji zniekształcenia w obszarze $\Omega_{\rho,n}$ jest jednocześnie rozwiązaniem lokalizacji zniekształcenia dla zadania (4) (czyli lokalizacji zniekształcenia w obszarze $\Omega_{\mathbf{D}}$), gdyż wartości własne dla obu zadań są takie same.

W zadaniu (4) jednej wartości własnej może odpowiadać kilka funkcji własnych, z których każda jest określona dla innej deformacji. Konstrukcja nowego zadania (12) oprócz zmiany operatora z A na $-A$ przyjmuje wartości własne jako parametry. Nie jest to zadanie spektralne. Jego rozwiązaniem są funkcje, które na fragmencie brzegu bliskie są odpowiednim funkcjom własnym, będącym rozwiązaniem zadania (4). W ten sposób zlokalizowane zniekształcenie będzie bliskie zniekształceniu obszaru $\Omega_{\mathbf{D}}$ z pewnym prawdopodobieństwem. Dla wyznaczenia wartości oczekiwanej lokalizacji zniekształcenia uwzględniamy wyniki otrzymane dla każdej wartości λ_i , $i = 1, 2, \dots, r$.

4. Lokalizacja zniekształcenia obszaru $\Omega_{\mathbf{D}}$. Opiszmy wcześniejsze rozważania w terminach teorii prawdopodobieństwa.

Niech $\Xi = (\Omega, \mathbf{F}, P)$ będzie zupełną przestrzenią probabilistyczną ($\Omega = [0, 1] \times [0, 1]$, \mathbf{F} – σ ciało podzbiorów Ω , P – miara prawdopodobieństwa). Niech \mathbf{D}_t , $t = 1, 2, \dots$ będzie ciągiem zdarzeń losowych określonych w przestrzeni probabilistycznej Ξ . Dla każdego $t = 1, 2, \dots$ zdarzenie losowe polega na wygenerowaniu zniekształcenia \mathbf{D}_t w obszarze $[0, 1] \times [0, 1]$. Zakładamy, że wygenerowane zniekształcenie spełnia warunki założone wcześniej, to znaczy jest zawarte we wnętrzu obszaru $[0, 1] \times [0, 1]$ oraz jego wielkość (rozumiana tutaj jako pole powierzchni) jest bardzo mała – bliska zeru. Stosując wcześniejsze oznaczenia, obszar $[0, 1] \times [0, 1]$ z wygenerowanym zniekształceniem \mathbf{D}_t oznaczamy $\Omega_{\mathbf{D}_t}$.

Niech X_t, Y_t $t = 1, 2, \dots$ będą ciągami zmiennych losowych odpowiadających ciągowi zdarzeń losowych \mathbf{D}_t , $t = 1, 2, \dots$. Ciągi te określone są następująco. X_t , $t = 1, 2, \dots$ jest ciągiem zmiennych losowych, którego składowe określają r ustalonych wartości własnych, będących rozwiązaniami (4) zdefiniowanego dla deformacji \mathbf{D}_t , $t = 1, 2, \dots$. Generujemy ciąg zmiennych losowych, co oznacza, że każdy element ciągu X_t , $t = 1, 2, \dots$ odpowiada innemu problemowi postaci (4). Innymi słowy, dla każdego t rozważamy nowe zadanie w nowym obszarze $\Omega_{\mathbf{D}_t}$ z innym zniekształceniem \mathbf{D}_t . Każde nowe zadanie jest postaci (4), ale odpowiada innemu zniekształceniu rozważa-

nego obszaru, to znaczy innemu obszarowi $\Omega_{\mathbf{D}_t}$. Y_t , $t = 1, 2, \dots$ jest ciągiem dwuwymiarowych zmiennych losowych odpowiadających ciągowi zdarzeń losowych \mathbf{D}_t , $t = 1, 2, \dots$. Dla każdego $t = 1, 2, \dots$, Y_t jest wektorem dwuelementowym określającym położenie zniekształcenia \mathbf{D}_t w obszarze $\Omega_{\mathbf{D}_t}$. Przypomnijmy, że każdy Y_t jest wyliczany następująco: zniekształcenie \mathbf{D}_t jest pokryte L kołami $\mathbf{B}_l^t(\bar{y}_l, \rho)$, $l = 1, 2, \dots, L$, każde o promieniu ρ i środku \bar{y}_l . Każdemu z kół przypisujemy liczby p_l^1, p_l^2 , $l = 1, 2, \dots, L$ takie, że dla każdego $l = 1, 2, \dots, L$, $p_l^1 > 0$, $p_l^2 > 0$ oraz $\sum_{l=1}^L p_l^1 = 1$, $\sum_{l=1}^L p_l^2 = 1$. Ponadto wartości p_l^1 i p_l^2 są dobrane tak, że im większa ich część pokrywa się ze zniekształceniem tym wyższe są wartości p_l^1 oraz p_l^2 . Pokrycie jest zdefiniowane w taki sposób, że wszystkie koła mają ten sam promień, natomiast zniekształcenie \mathbf{D}_t składa się z fragmentów różnej szerokości. Zachodzi zależność $\mathbf{D}_t \subset \bigcup_{l=1}^L \mathbf{B}_l^t(\bar{y}_l, \rho)$ ponadto może zachodzić $\bigcup_{l=1}^L \mathbf{B}_l^t(\bar{y}_l, \rho) \cap \Omega_{\mathbf{D}_t} \neq \emptyset$. Tym samym mogą być takie koła z pokrycia zniekształcenia \mathbf{D}_t , których bardzo mała część pola pokrywa zniekształcenie natomiast pozostała część pokrywa fragment $\Omega_{\mathbf{D}_t}$ przyległy zniekształceniu \mathbf{D}_t . Dla takich kół liczby p_l^1, p_l^2 są bliskie 0. Y_t , $t = 1, 2, \dots$ są wyliczane na podstawie pokrycia w następujący sposób (wzór (9)) $Y_t = (Y_t^1, Y_t^2)$, $t = 1, 2, \dots$, gdzie $Y_t^1 = \sum_{l=1}^L (p_l^1 \cdot \bar{y}_l^1)$, $Y_t^2 = \sum_{l=1}^L (p_l^2 \cdot \bar{y}_l^2)$, $t = 1, 2, \dots$

Reasumując, ciągi zmiennych losowych X_t, Y_t , $t = 1, 2, \dots$ generowane są w następujący sposób. W obszarze $[0, 1] \times [0, 1]$ wprowadzamy zniekształcenie \mathbf{D}_t . Otrzymujemy w ten sposób obszar $\Omega_{\mathbf{D}_t}$. Definiujemy pokrycie zniekształcenia \mathbf{D}_t , to znaczy dla ustalonych L i ρ określamy środki kół \bar{y}_l , $l = 1, 2, \dots, L$ stanowiących pokrycie \mathbf{D}_t . Dla każdego z tych kół określamy liczby p_l^1 oraz p_l^2 , $l = 1, 2, \dots, L$. Na podstawie położenia kół generujących pokrycie zniekształcenia \mathbf{D}_t oraz przypisanych im liczb p_l^1, p_l^2 wyliczamy wektor Y_t . Następnie rozwiązujemy zadanie (4). Spośród wszystkich wartości własnych, będących jego rozwiązaniem wybieramy r , ustalonych, które definiują współrzędne wektora X_t .

Taki sposób obliczeń spowodowany jest zależnością pomiędzy wartościami X_t a Y_t , $t = 1, 2, \dots$. Po pierwsze, na początku znane jest zniekształcenie. Po drugie, wartości wektora X_t w sposób pośredni zależą od wartości wektora Y_t . Dzieje się tak dlatego, że ciągi zmiennych losowych X_t i Y_t , $t = 1, 2, \dots$ odpowiadają temu samemu ciągowi zdarzeń losowych \mathbf{D}_t , $t = 1, 2, \dots$ i każdy z nich w inny sposób opisuje ten sam ciąg zdarzeń losowych. Wektor Y_t stanowi swego rodzaju opis zniekształcenia \mathbf{D}_t przez to że określa jej lokalizację, natomiast wartości współrzędnych wektora X_t opisują własności zniekształcenia jako rozwiązania zależnego od deformacji problemu spektralnego (4). Naturalną jest taka kolejność obliczeń, w której najpierw wyliczane są wartości Y_t , a następnie wartości X_t . X_t jest jednym z rozwiązań zadania postaci (4), ale aby je rozwiązać musimy najpierw znać

położenie, którego opisem jest wektor Y_t i wielkość deformacji.

Z praktycznego punktu widzenia interesuje nas przypadek odwrotny. Chcemy na podstawie wartości wektora X_t (r ustalonych wartości własnych) lokalizować zniekształcenie obszaru. Wartości własne, z których r ustalonych to składowe wektora X_t są pewnymi wartościami fizycznymi. Na podstawie X_t chcemy określić położenie zniekształcenia. Nie będzie to precyzyjny opis zniekształcenia, a jedynie przybliżona, oczekiwana jego lokalizacja. Definicja wektora Y_t (9) jest sformułowana tak, że określa przybliżone miejsce położenia zniekształcenia wskazując nie cały jej obszar, a jedynie jeden punkt. Rozważany problem odwrotny można zdefiniować jako szukanie wektora Y_t na podstawie X_t .

Twierdzenie 2 pokazuje, że na podstawie jednej, ustalonej wartości własnej można jednoznacznie określić położenie zniekształcenia. Jest to możliwe dla z góry zadanego obszaru $\Omega_{\rho,n}$. Rozważany tutaj problem odwrotny jest rozumiany jako rozwiązywanie zadania mającego na celu określenie lokalizacji zniekształcenia obszaru $[0, 1] \times [0, 1]$ na podstawie z góry ustalonej wartości własnej. W tym przypadku wartości własne nie są wyliczane tylko pochodzą z obserwacji jakiegoś zjawiska. Nie jest znana postać deformacji, więc szukany wektor Y_t nie może być wyliczony na podstawie wzoru (9), ale musi być wyliczany w inny sposób.

Procedura wyliczania wektora Y_t jest złożona i polega na r -krotnym wyliczeniu w oparciu o Twierdzenie 2 lokalizacji zniekształcenia, za każdym razem uwzględniając inną z r ustalonych wartości własnych. Ponieważ dla jednej wartości własnej może istnieć kilka różnych funkcji własnych, więc parametrami wejściowymi do stosowania Twierdzenia 2 mogą być różne funkcje. Stąd dla wektora r wartości własnych wyliczone środki dziur na podstawie Twierdzenia 2 mogą być różne. Znając dodatkowe informacje, pomagające w określeniu fragmentu obszaru, w którym może znajdować się zniekształcenie, definiujemy rozkład prawdopodobieństwa. Rozkład ten odnosi się do otrzymanych r wyników i jest definiowany w ten sposób, że daje większe prawdopodobieństwo tym lokalizacjom zniekształcenia, które są w obszarze wskazanym przez dodatkowe informacje. Jednocześnie dla lokalizacji zniekształceń znajdujących się poza wskazanym obszarem, wartości prawdopodobieństwa są bardzo małe. Na podstawie zdefiniowanego rozkładu wyliczamy wartość oczekiwaną każdej współrzędnej określającej położenie zniekształcenia i otrzymujemy wektor, który traktujemy jako prawdopodobne położenie zniekształcenia odpowiadające wektorowi zadanych r wartości własnych.

4.1 Aproksymacja odwzorowania odwrotnego. W poprzedniej sekcji zdefiniowano zadanie odwrotne, które dla argumentu X_t zwróci szukaną wartość wektora Y_t . Odwzorowanie to nie jest funkcją. Jeżeli X_t jest argumentem

odwzorowania odwrotnego, to wobec braku jednoznaczności możemy mówić o zbiorze wartości odpowiadających argumentowi X_t , nie zaś o jednej, konkretnej wartości wektora Y_t . Z tego powodu nie daje się przybliżyć tak zdefiniowanego odwzorowania odwrotnego. Zamiast niego rozważamy inne odwzorowanie, które jest funkcją i ma na celu dla ustalonego X_t wyznaczyć wartość oczekiwaną lokalizacji zniekształcenia D_t .

Niech dane będzie odwzorowanie θ_0 , które jest warunkową wartością oczekiwaną zdefiniowaną następująco

$$(17) \quad \theta_0(X_t) = E(Y_t | X_t),$$

gdzie $t = 1, 2, \dots$. Jest to odwzorowanie, które zmiennej losowej X_t przyporządkowuje wektor warunkowych wartości oczekiwanych zmiennej losowej Y_t . $E(Y_t | X_t) = [E(Y_t^1 | X_t), E(Y_t^2 | X_t)]$, gdzie $Y_t = [Y_t^1, Y_t^2]$, zaś $E(Y_t^1 | X_t)$ oraz $E(Y_t^2 | X_t)$ są warunkowymi wartościami oczekiwanymi kolejnych współrzędnych wektora Y_t przy założeniu, że X_t jest dane. Odwzorowanie θ_0 zwraca zatem wartość oczekiwaną położenia zniekształcenia D_t w obszarze Ω_{D_t} dla danej wartości własnej X_t . Naszym celem jest aproksymacja odwzorowania θ_0 .

Niech dany będzie ciąg

$$(18) \quad Z_t = (Y_t^T, X_t^T)^T,$$

gdzie $t = 1, 2, \dots, T$ – oznacza transpozycję wektorów. Elementy ciągu Z_t zdefiniowane są za pomocą wartości ciągów X_t i Y_t , $t = 1, 2, \dots$. X_t jest ustaloną wartością własną problemu (4), zaś Y_t to wektor określający położenie zniekształcenia D_t . Jeżeli dla wszystkich możliwych wektorów X_t znamy elementy ciągu Z_t to mamy pełną i dokładną informację o odwzorowaniu θ_0 . W praktyce taka sytuacja nie zachodzi, to znaczy znamy jedynie przybliżone wartości odwzorowania θ_0 dla skończonej liczby argumentów. Naszym celem jest aproksymacja, czyli przybliżenie nieznanego odwzorowania θ_0 na podstawie posiadanej częściowej, przybliżonej informacji. Oznacza to, że wyznaczamy nowe odwzorowanie, którego wartości będą bliskie wartościom odwzorowania θ_0 dla argumentów X_t .

Dla celów praktycznych ograniczamy się do skończonej liczby zdarzeń losowych, które tworzą ciąg Z_t . Niech N oznacza liczebność zbioru $Z^N = (Z_1, Z_2, \dots, Z_N)$. Procedura aproksymująca jest realizowana dla zbiorów Z^N , które tworzą tak zwany zbiór uczący. Odwzorowanie θ_0 jest odwzorowaniem zdefiniowanym dla zmiennych losowych. Tym samym jest elementem probabilistycznej przestrzeni odwzorowań. Oznaczmy tę przestrzeń przez Θ .

Niech N będzie z góry ustaloną liczbą. Dla każdego (ustalonego) N prowadzimy procedurę aproksymacji odwzorowania θ_0 . Przez procedurę aproksymacji rozumiemy tutaj mechanizm tworzenia nowego odwzorowania, które jest zdefiniowane dla tych samych przestrzeni co odwzorowanie θ_0 oraz

dla argumentów X_t pochodzących ze zbioru Z^N przyjmuje wartości bliskie $E(Y_t | X_t)$. Procedura ta jest powtarzana dla coraz większych wartości N . To oznacza, że otrzymujemy kolejne odwzorowania, których wartości są coraz bliższe wartościom odwzorowania θ_0 dla coraz większej liczby argumentów. Przyjmując jako indeks liczbę elementów zbioru uczącego otrzymujemy ciąg, który aproksymuje odwzorowanie θ_0 . Przedstawiona konstrukcja umożliwia otrzymanie zbieżności aproksymacji nieznanego odwzorowania θ_0 przy $N \rightarrow \infty$.

Niech $\hat{\theta}_N$, $N = 1, 2, \dots$ będzie ciągiem odwzorowań aproksymujących odwzorowanie θ_0 takim, że dla każdego N , $\hat{\theta}_N$ należy do przestrzeni Θ . $\hat{\theta}_N$, $N = 1, 2, \dots$ jest procesem stochastycznym (patrz [3] strona 286), gdyż każde z odwzorowań $\hat{\theta}_N$ jest odwzorowaniem z przestrzeni probabilistycznej.

Stworzony ciąg aproksymujący to estymator (patrz [3] strona 479). Minimalnym założeniem procedury aproksymującej jest zgodność estymatora (patrz [3] strona 479). Przez zgodność estymatora rozumiemy, że błąd aproksymacji odwzorowania θ_0 ciągiem $\hat{\theta}_N$, $N = 1, 2, \dots$ dąży do zera przy wielkości zbioru uczącego dążącego do nieskończoności względem miary prawdopodobieństwa. Innymi słowy

$$\lim_{N \rightarrow \infty} P(|\hat{\theta}_N - \theta_0| \geq \varepsilon) = 0,$$

gdzie ε jest dowolnie małą, z góry ustaloną liczbą, zaś $P(a)$ oznacza prawdopodobieństwo, że zdarzenie a zachodzi.

Błędy końcowe stosowanej procedury dla każdego ustalonego N są nieuniknione i powstają na skutek przybliżonych wartości X_t , Y_t oraz braku pełnej informacji o funkcji θ_0 . Z uwagi na wzrost wielkości zbioru uczącego błędy procedury aproksymacji maleją przy $N \rightarrow \infty$ co umożliwia zbieżność metody.

5. Sieć Elmana jako aproksymator. W celu przeprowadzenia procedury aproksymacji odwzorowania θ_0 stosujemy sieć neuronową typu Elmana. Wybór tego typu sieci nie jest przypadkowy. Dane uczące, którymi dysponujemy, nie są danymi dokładnymi, a jedynie stanowią ich pewne przybliżenie. W sieci typu Elmana występuje sprzężenie zwrotne, które sprawia, że zapamiętywane są nie tylko pojedyncze wzorce, ale również uwzględniane są relacje między nimi. Dzieje się tak dlatego, że w kolejnych iteracjach część współrzędnych wektora wejściowego stanowią wartości wektora ze zbioru uczącego oraz część współrzędnych wektora wejściowego, to wartości wyjściowe neuronów pierwszej warstwy z poprzedniej iteracji. Taki obieg informacji w sieci pozwala na bardziej efektywne wykorzystanie elementów zbioru uczącego jako całości. Ponadto sieć typu Elmana uczona na podstawie przybliżonych danych z większą skutecznością aproksymuje szukane odwzorowanie (por. [7]).

6. Idea aproksymacji – zgodność „sit” estymatorów. Idea aproksymacji opisana w niniejszej pracy jest adaptacją idei aproksymacji zaproponowanej w pracy [14]. W przypadku przedstawionym w niniejszej pracy idea ta odnosi się do sieci Elmana działającej z \mathbb{R}^r do \mathbb{R}^2 .

W celu omówienia procedury aproksymacji szukanego odwzorowania i udowodnienia jej zgodności przedstawimy najpierw krótki opis całości stosowanej metody.

Do realizowania procedury aproksymacji stosujemy sieć neuronową Elmana, która składa się z r wejść, wzbogaconych o stały sygnał bias oraz z q wejść kontekstowych. Ponadto sieć składa się z dwóch warstw, pierwsza zawiera q neuronów, zaś druga – dwa. Dla pierwszej warstwy stosuje się funkcję aktywacji ψ , zaś dla warstwy drugiej stosuje się liniową funkcję aktywacji. W praktyce oznacza to, że wyjście z sieci jest równe wartości sumatorów neuronów warstwy wyjściowej (drugiej). Wagi warstwy pierwszej są oznaczane przez γ_{ji} , $i = 0, \dots, q+r$, $j = 1, \dots, q$, zaś wagi warstwy drugiej są oznaczane przez β_{ji} , $i = 0, 1, \dots, q$, $j = 1, 2$.

Niech δ^q oznacza zbiór wszystkich wag sieci. Wówczas możemy zdefiniować odwzorowanie $f^q(\cdot, \delta^q) : \mathbb{R}^r \rightarrow \mathbb{R}^2$ takie, że

$$f^q(x, \delta^q) = (y_1^2, y_2^2),$$

gdzie $x = [1, x_1^1, x_2^1, \dots, x_{q+r}^1]$, zaś

$$y_1^2 = \beta_{10} + \sum_{j=1}^q \beta_{1j} \psi \left(\sum_{i=0}^{q+r} x_i^1 \gamma_{ji} \right),$$

oraz

$$y_2^2 = \beta_{20} + \sum_{j=1}^q \beta_{2j} \psi \left(\sum_{i=0}^{q+r} x_i^1 \gamma_{ji} \right).$$

Konstruujemy ciąg aproksymatorów dla θ_0 przez konstruowanie sieci Elmana składających się z q neuronów w warstwie ukrytej, gdzie wartość q będzie wzrastać w odpowiedniej relacji do N . N określa liczbę elementów, z których składa się zbiór uczący. Tworzenie ciągu aproksymatorów polega na tym, że dla wybranego ciągu wag $\hat{\delta}_N$ tworzymy ciąg sieci $\hat{\theta}_N \equiv f^{q_n}(\cdot, \hat{\delta}_N)$, który jest aproksymatorem nieznannej funkcji θ_0 .

Dla bardziej precyzyjnego zdefiniowania rozwiązania dla zadania polegającego na szukaniu ciągu $\hat{\theta}_N$ definiujemy

$$T(\psi, q, \Delta) = \left\{ \theta \in \Theta : \theta(\cdot) = f^q(\cdot, \delta^q), \sum_{i=0}^q |\beta_{ji}| < \Delta, j = 1, 2, \sum_{j=1}^q \sum_{i=0}^{q+r} |\gamma_{ji}| < q\Delta \right\}$$

zbiór wszystkich funkcji, którymi są sieci neuronowe o konstrukcji przedstawionej powyżej. Ponadto konstruujemy ciąg sit $\{\Theta_N(\psi)\}$ za pomocą ciągów $\{q_N\}$ i $\{\Delta_N\}$ zdefiniowany jako $\Theta_N(\psi) = T(\psi, q_N, \Delta_N)$, gdzie $N = 1, 2, \dots$. Dla ciągów $\{q_N\}$ i $\{\Delta_N\}$ estymator sit $\hat{\theta}_N$ jest zdefiniowany jako rozwiązanie problemu najmniejszych kwadratów

$$(19) \quad \min_{\theta \in \Theta_N(\psi)} \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \theta(X_t)]^2,$$

$N = 1, 2, \dots$

Rozwiązaniem problemu (19) jest estymator $\hat{\theta}_N$, dla którego $\hat{\delta}_N$ oznacza wszystkie wagi sieci. W pracy określimy warunki, jakie należy nałożyć na $\{Z_t\}$, ψ , $\{q_N\}$ oraz $\{\Delta_N\}$, aby zapewnić zgodność estymatora $\hat{\theta}_N$ z θ_0 . Przed określeniem tych warunków należy podkreślić, że chcielibyśmy otrzymać globalne rozwiązanie zadania (19). Wiadomo, że klasycznie stosowane metody takie jak wsteczna propagacja błędów czy metoda iteracyjna Newtona dostarczają jedynie ekstrema lokalne dlatego w osiągnięciu naszego celu są nieskuteczne. Wymagane jest więc zastosowanie w ich miejsce globalnych metod optymalizacyjnych opisanych na przykład w pracach [1] i [15]. W tych pracach zawarta jest istota informacji na temat tego typu optymalizacji.

Bez zmniejszania ogólności rozważań założymy, że globalne ekstremum zadania (19) jest osiągnięte. W pracy zostaną podane warunki, jakie muszą zostać spełnione, aby założenie o tym, że globalne rozwiązanie jest osiągnięte, było spełnione.

Założymy, że $\{Z_t\}$ jest procesem stochastycznym, ponadto założymy, że θ_0 jest funkcją całkowalną określoną na podzbiórze ograniczonym przestrzeni \mathbb{R}^r . Przyjmijmy, że ψ jest funkcją aktywacji, która spełnia warunek Lipschitza ze stałą L .

Omawiając tę metodę należy zauważyć, że stosowana w niej sieć zmienia się w czasie realizacji procedury. Zmiana ta polega na tym, że wraz ze wzrostem liczebności zbioru uczącego N wzrasta liczba neuronów w warstwie ukrytej (i tym samym w warstwie kontekstowej). Procesowi temu przypisana jest odpowiednia relacja w jakiej mają się zmienić ograniczenia na wagi sieci w stosunku do rosnącego parametru N . Muszą być spełnione następujące warunki: $\Delta_N \rightarrow \infty$ przy $N \rightarrow \infty$ oraz $\Delta_N = o(N^{1/4})$, co oznacza, że $n^{-1/4}\Delta_N \rightarrow \infty$ przy $N \rightarrow \infty$. Ponadto odpowiedni wybór $\{q_N\}$ jest zależny od $\{\Delta_N\}$ oraz własności $\{Z_t\}$. W przypadku, gdy $\{Z_t\}$ jest procesem stochastycznym o niezależnym, jednakowym rozkładzie wówczas $q_N \rightarrow \infty$ przy $N \rightarrow \infty$ oraz $q_N \Delta_N^4 \log q_N \Delta_N = o(N)$. Jeżeli $\{Z_t\}$ jest mieszanym procesem stochastycznym wówczas $q_N \rightarrow \infty$ przy $N \rightarrow \infty$ oraz $q_N \Delta_N^2 \log q_N \Delta_N = o(N^{1/2})$.

Przyjęta zasada wzrostu wielkości warstwy ukrytej wraz ze wzrostem

liczebności zbioru uczącego sprawia, że metoda aproksymacji staje się bardziej skuteczna. Ponadto należy zauważyć, że wprowadzone wzajemne relacje między parametrami sieci, a wzrostem wartości N pokazują, że wartość Δ_N wzrasta w wolnym tempie w stosunku do N , zaś wartość q_N wzrasta w szybkim tempie w stosunku do N .

Uwzględniając definicję sieci neuronowej, jako funkcji na którą nałożone są pewne ograniczenia, możemy zapisać następujące zależności $\sum_{i=0}^q |\beta_{1i}| < \Delta_N$, $\sum_{i=0}^q |\beta_{2i}| < \Delta_N$, $\sum_{j=1}^q \sum_{i=0}^{q+r} |\gamma_{ji}| < q_N \Delta_N$, czyli $\sum_{i=0}^{q+r} |\gamma_{ji}| < \Delta_N$ dla każdego $j = 1, 2, \dots, q$.

Pomimo wprowadzenia pewnych warunków na tempo wzrostu parametrów sieci – ograniczeń na wagi, nie istnieją praktyczne metody łatwego doboru wielkości sieci w stosunku do danego N . Jedną ze stosowanych metod nazywana jest metodą krzyżowanej poprawności walidacji (ang. *cross-validated complexity*) i została omówiona w pracy [12]. Metoda ta polega na tym, że zakładamy, że $\{\Delta_N\}$ jest dane. Kładziemy $\Theta_N(\psi, q) = T(\psi, q, \Delta_N)$. Dla danego N , sieć składająca się z q neuronów warstwy ukrytej oraz ma wyliczony średniokwadratowy błąd na zbiorze uczącym zgodnie z następującą formułą

$$\min_{\theta \in \Theta_N(\psi, \delta)} \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \theta(X_t)]^2.$$

Niestety, ta miara jest zbyt optymistyczna. Wartość błędu maleje, gdy q wzrasta. Dla obserwacji $(Y_t^T, X_t^T)^T$, wartość $\hat{\theta}_N^q$ jest rozwiązaniem powyższego problemu minimalizacji. To sprawia, że wartość błędu kwadratowego t -tej obserwacji $([E(Y_t | X_t) - \hat{\theta}_N^q(X_t)]^2)$ jest zbyt mały. Istotnie, wielkość sieci dobrana za pomocą q generuje błąd średniokwadratowy tak mały jak to tylko możliwe; błąd kwadratowy może być równy 0 przy wyborze $q = N$.

Metoda krzyżowej poprawności unika tego typu trudności. Błąd średniokwadratowy jest estymowany przez błąd kwadratowy każdej obserwacji z wyjątkiem t -tej obserwacji. Bardziej formalnie, metoda krzyżowej poprawności definiuje

$$C_N(q) \equiv \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \hat{\theta}_{N(t)}^q(X_t)]^2.$$

W pełni automatyczna metoda ustala liczbę neuronów warstwy ukrytej na wartość \hat{q}_N , która jest najmniejszym rozwiązaniem zadania

$$\min_{q \in N_N} C_N(q),$$

gdzie N_N jest zbiorem N pierwszych liczb naturalnych, to znaczy $N_N = \{1, 2, \dots, N\}$. Wartość \hat{q}_N będziemy nazywać złożonością metody krzyżowej poprawności. Wówczas estymatorem sit metody krzyżowej poprawności

będziemy rozumieć rozwiązanie $\tilde{\theta}_N$ problemu

$$\min_{\theta \in \Theta_N(\psi, \hat{q}_N)} \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \theta(X_t)]^2.$$

Podsumowując, sieć Elmana może być uczona aproksymacji nieznanej funkcji w rozumieniu procedury zgodności omówionej powyżej. W celu otrzymania zgodności aproksymacji konieczne jest aby sieć zmieniała swoją strukturę podczas trwania procesu uczenia. Przez zmianę struktury należy rozumieć wzrost liczby neuronów w warstwie ukrytej wraz ze wzrostem wielkości zbioru uczącego. Ponadto wymagane jest też spełnienie warunków nałożonych na ograniczenia wag, których wartość wzrasta wraz ze wzrostem N , jednak w ściśle określonej relacji.

7. Sieć neuronowa typu Elmana jako aproksymator warunkowej wartości oczekiwanej. W celu aproksymowania odwzorowania θ_0 konstruujemy sieć neuronową typu Elmana, w której wprowadzamy pewną modyfikację. Modyfikacja ta dotyczy wejść kontekstowych do sieci. W ogólnym przypadku wejścia te przyjmują wartości wyjściowe neuronów warstwy pierwszej otrzymane w poprzedniej iteracji (dla pierwszej iteracji przyjmują wartość zero). W przypadku rozważanym w pracy wprowadzamy pewne ograniczenie na te wartości przyjmując, że każde wejście kontekstowe będzie miało wartość równą odpowiadającemu mu wyjściu neuronu warstwy pierwszej z poprzedniej iteracji podzielone przez liczbę neuronów warstwy ukrytej. Jest to konieczne ze względu na stosowaną metodę aproksymacji. W rozważanym przypadku aproksymator – czyli sieć neuronowa, zwiększa liczbę neuronów w warstwie ukrytej w miarę postępu procesu aproksymacji. Z uwagi na to, że liczba wejść kontekstowych jest równa liczbie neuronów w warstwie ukrytej, jednocześnie zwiększa się wymiar wektora wejściowego do sieci. Jest więc konieczne ograniczenie wzrostu wartości sumatora neuronów warstwy ukrytej. Brak takiego ograniczenia powoduje, że wraz ze wzrostem liczby wejść kontekstowych maleje udział właściwego wektora wejściowego w wartości sumatora każdego z neuronów warstwy ukrytej. Wprowadzone ograniczenie powoduje, że suma wartości wszystkich wejść kontekstowych nie przekracza jeden niezależnie od ich liczby. Jest to spowodowane tym, że każde wyjście z neuronu warstwy pierwszej jest z zakresu $(0, 1)$, a podzielone przez liczbę neuronów q i zsumowane po q wejściach daje nie więcej niż 1.

Uwzględniając powyższe założenia stosowana w pracy sieć ma r wejść, którymi są składowe wektora X_t oraz wejścia dodatkowe: bias i neurony kontaktowe. Pierwsza warstwa sieci składa się z q neuronów, każdy z funkcją aktywacji ψ . Warstwa ostatnia składa się z dwóch neuronów, każdy z liniową

funkcją aktywacji⁽²⁾.

Każde połączenie w sieci (za wyjątkiem połączeń neuronów pierwszej warstwy z neuronami kontekstowymi) posiada wagę. Niech γ_{ji} , $i = 0, \dots, q+r$, $j = 1, \dots, q$ będą wagami połączeń warstwy pierwszej zdefiniowanymi w ten sposób, że γ_{ji} odpowiada połączeniu z i -tego wejścia sieci do j -tego neuronu pierwszej warstwy. Niech β_{ji} , $i = 0, 1, \dots, q$, $j = 1, 2$ będą wagami połączeń warstwy drugiej określonymi w ten sposób, że β_{ji} odpowiada połączeniu z i -tego neuronu pierwszej warstwy z j -tym neuronem warstwy drugiej.

Niech $[1, x_1^1, x_2^1, \dots, x_{q+r}^1]$ będzie wektorem wejściowym do sieci. Jest to zapis wektora wejściowego, który zawiera bias 1 oraz wejścia kontekstowe $x_{r+1}^1(t) = y_1^1(t-1)/q$, \dots , $x_{r+q}^1 = y_q^1(t-1)/q$. Wartości x_1^1, \dots, x_r^1 są współrzędnymi wektora wejściowego, którego wartości stanowią elementy wejściowe do sieci. Niech $[y_1^2, y_2^2]$ będzie wektorem wyjściowym z sieci. Niech ponadto ψ oznacza funkcję aktywacji neuronów warstwy pierwszej. Wówczas informacja wyjściowa z sieci wyliczana jest następująco

$$y_1^2 = \beta_{10} + \sum_{j=1}^q \beta_{1j} \psi \left(\sum_{i=0}^{q+r} x_i^1 \gamma_{ji} \right),$$

oraz

$$y_2^2 = \beta_{20} + \sum_{j=1}^q \beta_{2j} \psi \left(\sum_{i=0}^{q+r} x_i^1 \gamma_{ji} \right).$$

Niech δ^q oznacza zbiór wszystkich wag sieci. Wówczas możemy zdefiniować odwzorowanie $f^q(\cdot, \delta^q) : \mathbb{R}^r \rightarrow \mathbb{R}^2$ takie, że

$$f^q(x, \delta^q) = (y_1^2, y_2^2),$$

gdzie $x = [1, x_1^1, x_2^1, \dots, x_{q+r}^1]$, zaś y_1^2, y_2^2 zdefiniowane powyżej.

Liczba neuronów warstwy ukrytej wynosi q i w miarę realizacji procedury aproksymującej odwzorowanie θ_o będzie wzrastać do nieskończoności, czyli będzie stanowić parametr modelu. Innymi słowy, w celu aproksymacji odwzorowania θ_o konstruujemy ciąg aproksymatorów, którymi są sieci neuronowe, w których liczba neuronów w warstwie ukrytej wzrasta do nieskończoności wraz ze wzrostem liczby elementów w zbiorze uczącym (N).

Konstruujemy ciąg aproksymujący θ_0 biorąc sieci, składające się z q neuronów ukrytych i q neuronów kontekstowych, gdzie q rośnie wraz ze wzrostem N w odpowiedniej relacji. Dla ustalonego N generujemy wagi $\hat{\delta}_N$ takie, że $\hat{\theta}_N = f^{qN}(\cdot, \hat{\delta}_N)$ stanowi najlepszą ze wszystkich dla ustalonego N aproksymację nieznanego odwzorowania θ_0 . Innymi słowy szukając rozwiązania

⁽²⁾ liniowa funkcja aktywacji stosowana w pracy to funkcja $f(x) = x$

$\hat{\theta}_N$ dla ustalonego N definiujemy

$$T(\psi, q, \Delta) = \left\{ \theta \in \Theta : \theta(\cdot) = f^q(\cdot, \delta^q), \sum_{i=0}^q |\beta_{ji}| < \Delta, j = 1, 2, \sum_{j=1}^q \sum_{i=0}^{q+r} |\gamma_{ji}| < q\Delta \right\}$$

jako zbiór wszystkich sieci neuronowych typu Elmana z jedną warstwą ukrytą oraz liczbą neuronów w warstwie ukrytej równą ilości neuronów kontekstowych równą q , z funkcją aktywacji dla neuronów warstwy ukrytej ψ oraz wagami spełniającymi określone warunki ograniczające definiowane przez Δ . Konstruujemy tak zwany ciąg sit, czyli $\{\Theta_N(\psi)\}$ przez zdefiniowanie ciągów $\{q_N\}$ i $\{\Delta_N\}$, kładąc $\Theta_N(\psi) = T(\psi, q_N, \Delta_N)$, $N = 1, 2, \dots$

Dla ustalonego N oraz zdefiniowanych ciągów $\{q_N\}$ i $\{\Delta_N\}$ aproksymacja sitami odwzorowania $\hat{\theta}_0$ jest zdefiniowana jako rozwiązanie następującego problemu najmniejszych kwadratów

$$(20) \quad \min_{\theta \in \Theta_N(\psi)} \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \theta(X_t)]^2, \quad N = 1, 2, \dots,$$

gdzie $E(Y_t | X_t)$ jest warunkową wartością oczekiwaną definiującą θ_0 dane wzorem (17).

Przedstawimy teraz podstawowe założenia o procesach stochastycznych oraz mierze prawdopodobieństwa, które są stosowane do konstrukcji procedury aproksymacji oraz dowodu jej zbieżności.

Przyjmujemy, zgodnie z wcześniejszymi oznaczeniami, że (Ω, \mathbf{F}, P) jest zupełną przestrzenią probabilistyczną, w której $\Omega = [0, 1] \times [0, 1]$, \mathbf{F} – σ ciało podzbiorów Ω oraz P – miara prawdopodobieństwa.

ZAŁOŻENIE 1. Niech Z_t , $t = 1, 2, \dots$ będzie realizacją procesu stochastycznego $\{Z_t : \Omega \rightarrow \mathbb{R}^{r+2}, t = 1, 2, \dots\}$ zdefiniowanego wzorem (18), określonego na (Ω, \mathbf{F}, P) oraz P będzie miarą prawdopodobieństwa taką, że proces stochastyczny $\{Z_t\}$ spełnia jeden z warunków

1. $\{Z_t\}$ jest procesem o niezależnym, jednakowym, rozkładzie albo
2. $\{Z_t\}$ jest stacjonarnym procesem mieszanym, spełniającym jeden z warunków: $\phi(k) = \phi_0 \rho_0^k$ albo $\alpha(k) = \alpha_0 \rho_0^k$, $k \geq 1$, dla pewnych stałych ϕ_0 , $\alpha_0 > 0$, $0 < \rho_0 < 1$, gdzie

$$\phi(k) = \sup_t \sup_{\{A \in \mathbf{F}_1^t, B \in \mathbf{F}_{t+k}^\infty : P(A) > 0\}} |P(B | A) - P(B)|,$$

$$\alpha(k) = \sup_t \sup_{\{A \in \mathbf{F}_1^t, B \in \mathbf{F}_{t+k}^\infty\}} |P(A \cap B) - P(A)P(B)|,$$

gdzie $\mathbf{F}_1^t \equiv \sigma(Z_1, \dots, Z_t)$, jest σ – ciałem wygenerowanym na (Z_1, \dots, Z_t) oraz $\mathbf{F}_t^\infty \equiv \sigma(Z_t, \dots)$, jest σ – ciałem wygenerowanym na (Z_t, Z_{t+1}, \dots) .

Położmy $\Theta = L^2(\mathbb{R}^r, \mu; \mathbb{R}^2)$, gdzie ρ , jest normą w tej przestrzeni. Zakładamy, że $\theta_0 = E(Y_t | X_t)$ jest elementem przestrzeni Θ .

DEFINICJA 1. Funkcję ψ nazywamy funkcją ściągającą, jeżeli $\psi : \mathbb{R} \rightarrow [0, 1]$, $\psi(a) \rightarrow 0$, gdy $a \rightarrow -\infty$, $\psi(a) \rightarrow 1$, gdy $a \rightarrow \infty$ oraz ψ jest monotoniczna.

DEFINICJA 2. Funkcję ψ nazywamy l -skończoną, jeżeli $\psi : \mathbb{R} \rightarrow \mathbb{R}$, jest l -krotnie różniczkowalna w sposób ciągły ($0 \leq l < \infty$) oraz

$$0 < \int_{\mathbb{R}} |D^l \psi(x)| dx < \infty,$$

gdzie $D^l \psi(x)$ oznacza l -tą pochodną funkcji ψ .

DEFINICJA 3. Mówimy, że $f(n) = o(g(n))$ wtedy i tylko wtedy, gdy dla każdego $c > 0$ istnieje $n_0 \in \mathbb{N}$, że dla każdego $n > n_0$ spełniony jest następujący warunek

$$0 < f(n) \leq cg(n).$$

ZAŁOŻENIE 2. $\Theta_N(\psi) = T(\psi, q_N, \Delta_N)$, gdzie $\psi : \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją aktywacji, która jest ograniczona, spełnia warunek Lipschitza oraz jest albo funkcją ściągającą albo l -skończoną dla $0 \leq l < \infty$. Ponadto $\{q_N\}$ i $\{\Delta_N\}$ są ciągami rosnącymi wraz ze wzrostem N w taki sposób, że $q_N \rightarrow \infty$, $\Delta_N \rightarrow \infty$, przy $N \rightarrow \infty$ oraz $\Delta_N = o(N^{\frac{1}{4}})$ i spełniony jest jeden z warunków

1. $q_N \Delta_N^4 \log q_N \Delta_N = o(N)$ albo
2. $q_N \Delta_N^2 \log q_N \Delta_N = o(N^{\frac{1}{2}})$.

8. Zbieżność aproksymacji. W celu udowodnienia zbieżności aproksymacji konieczne jest zastosowanie kilku lematów i twierdzeń.

LEMAT 1. Niech ciągi $\{q_N \in \mathbb{N}\}$ oraz $\{\Delta_N \in \mathbb{R}^+\}$ będą ciągami rosnącymi do nieskończoności przy N dążącym do nieskończoności. Jeżeli ψ jest funkcją ściągającą lub l -skończoną, dla dowolnego naturalnego l , $0 \leq l \leq \infty$, wówczas $\bigcup_{N=1}^{\infty} \Theta_N(\psi)$ jest gęsta w $\Theta = L^2(\mathbb{R}^r, \mu; \mathbb{R}^2)$.

TWIERDZENIE 3. Niech dla $N = 1, 2, \dots$, Θ_N - zdefiniowane wcześniej, będą zupełnymi, ośrodkowymi, borelowskimi podprzestrzeniami Θ . Niech $\hat{\Theta}_N : \Omega \rightarrow \Theta$ będzie takie, że wykres $gr \hat{\Theta}_N \in (\mathbf{F} \oplus \mathbf{B}(\Theta_N))$, gdzie $\mathbf{B}(\Theta_N)$ to σ ciało zbiorów borelowskich na Θ_N oraz że dla dowolnego $\omega \in \Omega$, $\hat{\Theta}_N(\omega)$ jest niepusty i zwarty oraz $\hat{\Theta}_N(\omega) \subset \Theta_N$.

Niech $Q_N : \Omega \times \Theta \rightarrow \mathbb{R}$ będzie $\mathbf{F} \oplus \mathbf{B}(\Theta)$ - mierzalne oraz załóżmy, że $Q_N(\omega, \cdot)$ jest półciągłe z dołu na Θ_N dla dowolnego ω w Ω , $N = 1, 2, \dots$

Wówczas dla każdego $N = 1, 2, \dots$ istnieje funkcja $\hat{\theta}_N : \Omega \rightarrow \Theta_N$ (stąd $\mathbf{F}/\mathbf{B}(\Theta_N)$ – mierzalna) taka, że

$$Q_N(\omega, \hat{\theta}_N(\omega)) = \min_{\theta \in \hat{\Theta}_N(\omega)} Q_N(\omega, \theta).$$

Dowód. To twierdzenie jest bezpośrednią konsekwencją rezultatów z paragrafu 2.c pracy [11], strony 501–503. \square

TWIERDZENIE 4. Załóżmy, że $\{\underline{\Theta}_N\}$ i $\{\bar{\Theta}_N\}$ są ciągami rosnących i zwartych podzbiorów Θ takich, że $\bigcup_{N=1}^{\infty} \underline{\Theta}_N$ jest gęsta w Θ i $\underline{\Theta}_N \subseteq \hat{\Theta}_N(\omega) \subseteq \bar{\Theta}_N$ dla wszystkich ω w Ω , $N = 1, 2, \dots$. Załóżmy, że istnieje funkcja $\bar{Q} : \Theta \rightarrow \bar{\mathbb{R}}$ taka, że dla wszystkich $\varepsilon > 0$

$$P[\omega : \sup_{\theta \in \bar{\Theta}_N} |Q_N(\omega, \theta) - \bar{Q}(\theta)| > \varepsilon] \rightarrow 0$$

przy $N \rightarrow \infty$ oraz dla $\theta_0 \in \Theta$

$$\inf_{\theta \in \eta^c(\theta_0, \varepsilon)} (\bar{Q}(\theta) - \bar{Q}(\theta_0)) > 0,$$

gdzie

$$\eta^c(\theta_0, \varepsilon) = \{\theta \in \Theta : \rho(\theta, \theta_0) \geq \varepsilon\}$$

i \bar{Q} jest ciągły w θ_0 . Wówczas $\rho(\hat{\theta}_N, \theta_0) \rightarrow 0$ względem prawdopodobieństwa P , co oznacza że dla dowolnego $\varepsilon > 0$, $P[\omega \in \Omega : \rho(\hat{\theta}_N(\omega), \theta_0) > \varepsilon] \rightarrow 0$ przy $N \rightarrow \infty$.

Dowód. Dowód jest analogiczny jak w pracy [14]. \square

DEFINICJA 4. Entropią metryczną domkniętego zbioru K nazywamy logarytm liczby zbiorów o promieniu ε , które stanowią pokrycie zbioru K i oznaczamy przez $H(\varepsilon)$.

TWIERDZENIE 5. Niech $\{\Theta_N\}$ będzie rosnącym ciągiem ośrodkowych podprzestrzeni Θ oraz $H_N(\varepsilon)$ będzie entropią metryczną Θ_N oraz połóżmy $G_N(\varepsilon) \equiv \exp H_N(\varepsilon)$.

Niech

$$\{s_N : \mathbb{R}^{r+2} \times \Theta_N \rightarrow \bar{\mathbb{R}}, N = 1, 2, \dots\}$$

oraz

$$\{m_N : \mathbb{R}^{r+2} \times \Theta_N \rightarrow \bar{\mathbb{R}}^+, N = 1, 2, \dots\}$$

będą ciągami funkcji takimi, że $s_N(\cdot, \theta)$ i $m_N(\cdot, \theta)$ są ciągłe na \mathbb{R}^{r+2} dla każdego θ z Θ_N . Załóżmy, że istnieje ciąg $\{d_N : \Theta_N \rightarrow \mathbb{R}^+\}$ i stała $\lambda > 0$ takie, że dla każdego $z \in \mathbb{R}^{r+2}$ i θ^0 z Θ_N

$$|s_N(z, \theta) - s_N(z, \theta^0)| < m_N(z, \theta^0) \rho(\theta, \theta^0)^\lambda$$

dla wszystkich θ z $\eta_N(\theta^0) \equiv \{\theta \in \Theta_N : \rho(\theta, \theta^0) < d_N(\theta^0)\}$.

Położmy

$$\begin{aligned}\bar{s}_N &\geq \sup_{z \in \mathbb{R}^{r+2}} \sup_{\theta \in \Theta_N} |s_N(z, \theta)|, \\ \bar{m}_N &\geq \sup_{z \in \mathbb{R}^{r+2}} \sup_{\theta \in \Theta_N} |m_N(z, \theta)|, \\ \underline{d}_N &\equiv \inf_{\theta \in \Theta_N} d_N(\theta)^\lambda,\end{aligned}$$

oraz załóżmy, że $\bar{m}_N \geq \underline{d}_N^{-1}$.

Niech $\{Z_t : \Omega \rightarrow \mathbb{R}^{r+2}\}$ będzie procesem stochastycznym na (Ω, \mathbf{F}, P) oraz $E(s_N(Z_t, \theta))$ oznacza wartość oczekiwaną $s_N(Z_t, \theta)$.

1. Jeżeli ciąg $\{Z_t\}$ jest procesem stochastycznym o niezależnym, jednakowym rozkładzie, wówczas dla dowolnego $\varepsilon > 0$ i dla dowolnego, odpowiednio dużego N

$$\begin{aligned}P \left[\sup_{\theta \in \Theta_N} \left| \frac{1}{N} \sum_{t=1}^N [s_N(Z_t, \theta) - E(s_N(Z_t, \theta))] \right| > \varepsilon \right] \\ \leq 2G_N \left(\left(\frac{\varepsilon}{6\bar{m}_N} \right)^{\frac{1}{\lambda}} \right) \left[\exp(-6N/7) + \exp(-\varepsilon^2 N / \bar{s}_{N,N}^2 [18 + 4\varepsilon]) \right].\end{aligned}$$

Jeśli dodatkowo ciąg $\{\Theta_N\}$ jest taki, że $N^{-1}(\bar{s})_N^2 \rightarrow 0$ przy $N \rightarrow \infty$ oraz dla dowolnego $\varepsilon > 0$

$$\left((\bar{s})_N^2 / N \right) H_N([\varepsilon / 6\bar{m}_N]^{\frac{1}{\lambda}}) \rightarrow 0$$

przy $N \rightarrow \infty$, wówczas dla dowolnego $\varepsilon > 0$

$$P \left[\sup_{\theta \in \Theta_N} \left| \frac{1}{N} \sum_{t=1}^N [s_N(Z_t, \theta) - E(s_N(Z_t, \theta))] \right| > \varepsilon \right] \rightarrow 0$$

przy $N \rightarrow \infty$.

2. Jeżeli ciąg $\{Z_t\}$ jest procesem stacjonarnym mieszanym spełniającym $\phi(k) = \phi_0 \rho_0^k$ albo $\alpha(k) = \alpha_0 \rho_0^k$, $\phi_0, \alpha_0 > 0$, $0 < \rho_0 < 1$, $k \geq 1$, wówczas istnieją stałe $0 < c_1, c_2 < \infty$ niezależne od N takie, że dla dowolnego $\varepsilon > 0$ i dowolnego N odpowiednio dużego

$$\begin{aligned}P \left[\sup_{\theta \in \Theta_N} \left| \frac{1}{N} \sum_{t=1}^N [s_N(Z_t, \theta) - E(s_N(Z_t, \theta))] \right| > \varepsilon \right] \\ \leq c_1 G_N([\varepsilon / 6\bar{m}_N]^{\frac{1}{\lambda}}) \left[\exp(-c_2 N^{\frac{1}{2}}) + \exp(-c_2 \varepsilon N^{\frac{1}{2}} / 6\bar{s}_N) \right].\end{aligned}$$

Jeśli dodatkowo ciąg $\{\Theta_N\}$ jest taki, że $N^{-1}(\bar{s})_N^{-2} \rightarrow 0$ przy $N \rightarrow \infty$ oraz dla dowolnego $\varepsilon > 0$

$$\left((\bar{s})_N^{-2} / N \right) H_N([\varepsilon / 6\bar{m}_N]^{\frac{1}{\lambda}}) \rightarrow 0$$

przy $N \rightarrow \infty$, wówczas dla dowolnego $\varepsilon > 0$

$$P \left[\sup_{\theta \in \Theta_N} \left| \frac{1}{N} \sum_{t=1}^N [s_N(Z_t, \theta) - E(s_N(Z_t, \theta))] \right| > \varepsilon \right] \rightarrow 0$$

przy $N \rightarrow \infty$.

Dowód. Dowód jest analogiczny jak w pracy [14]. \square

LEMAT 2. Niech dane będą: $\psi : \mathbb{R} \rightarrow \mathbb{R}$ – funkcją aktywacji, która jest ograniczona, spełnia warunek Lipschitza ze stałą L oraz jest albo funkcją ściągającą albo l – skończoną dla $0 \leq l < \infty$, $q \in \mathbb{N}$ oraz $\Delta \in \mathbb{R}^+$. Niech ρ_∞ będzie metryką określoną następująco: $\rho_\infty(\theta_1, \theta_2) \equiv \sup_{x \in \mathbb{R}^r} |\theta_1(x) - \theta_2(x)|$, gdzie $\theta_1, \theta_2 \in T(\psi, q, \Delta)$. Niech ponadto $H_\varepsilon(\psi, q, \Delta)$ oznacza entropię metryczną $T(\psi, q, \Delta)$ ze względu na ρ_∞ . Wówczas dla dowolnego $\varepsilon > 0$ odpowiednio małego

$$H_\varepsilon(\psi, q, \Delta) \leq p \log(8/\varepsilon) + p \log \left(\sqrt{2}(\Delta + L(2 + r\Lambda)\Delta^2) \right) + p \log p,$$

gdzie $p \equiv q(q + r + 3) + 2$.

LEMAT 3. Niech $\{q_N \in \mathbb{N}\}$ i $\{\Delta_N \in \mathbb{N}^+\}$ będą ciągami takimi, że $\{\Delta_N\}$ jest ciągiem rosnącym takim, że $\Delta_N \rightarrow \infty$, gdy $N \rightarrow \infty$. Połóżmy

$$\bar{s}_N = 4\Delta_N^2,$$

$$\bar{m}_N = 4\Delta_N,$$

$$\lambda = 1,$$

i dla ψ zdefiniowanej jak powyżej oraz $\varepsilon > 0$ połóżmy

$$H_N(\varepsilon) = H_\varepsilon(\psi, q_N, \Delta_N).$$

1. Jeśli

$$q_N \Delta_N^4 \log q_N \Delta_N = o(N),$$

wówczas dla dowolnego $\varepsilon > 0$

$$(\bar{s}_N^{-2}/N) H_N([\varepsilon/6\bar{m}_N]^{\frac{1}{\lambda}}) \rightarrow 0$$

przy $N \rightarrow \infty$;

2. Jeśli

$$q_N \Delta_N^2 \log q_N \Delta_N = o(N^{\frac{1}{2}}),$$

wówczas dla dowolnego $\varepsilon > 0$

$$(\bar{s}_N/N^{\frac{1}{2}}) H_N([\varepsilon/6\bar{m}_N]^{\frac{1}{\lambda}}) \rightarrow 0$$

przy $N \rightarrow \infty$.

Na podstawie powyższych lematów i twierdzeń można udowodnić twierdzenie, które pokazuje zgodność stosowanej aproksymacji odwzorowania θ_0 z $\hat{\theta}_N(X_t)$. Zgodność aproksymacji jest na poziomie ε . Ponadto pokazana

jest zbieżność stosowanej metody, która jest zbieżnością względem miary prawdopodobieństwa.

TWIERDZENIE 6. *Załóżmy, że spełnione są Założenia (1), (2), wówczas istnieje estymator $\hat{\theta}_N : \Omega \rightarrow \Theta$ taki, że*

$$\frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \hat{\theta}_N(X_t)]^2 = \min_{\theta \in \Theta_N(\psi)} \frac{1}{N} \sum_{t=1}^N [E(Y_t | X_t) - \theta(X_t)]^2,$$

$N = 1, 2, \dots$

Co więcej, dla dowolnego $\varepsilon > 0$, $P[\omega \in \Omega : \rho(\hat{\theta}_N(\omega), \theta_0) > \varepsilon] \rightarrow 0$ przy $N \rightarrow \infty$.

Dowód. Dowód jest analogiczny jak w pracy [14]. □

9. Podsumowanie. W niniejszej pracy przedstawiona została procedura przybliżonego wyznaczania lokalizacji uszkodzenia w zadanym obszarze. Rozważamy obszar $\Omega = [0, 1] \times [0, 1]$, w którym definiujemy problem spektralny (3). Rozwiązaniami tego problemu są wartości własne odpowiadające obszarowi Ω . Jest to zadanie wyjściowe dla prowadzonych rozważań.

W obszarze Ω wprowadzamy uszkodzenie (zniekształcenie) \mathbf{D} . Dla obszaru z deformacją $\Omega_{\mathbf{D}} = \Omega \setminus \mathbf{D}$ rozważamy problem spektralny (4), którego rozwiązaniem jest nieskończony ciąg par (λ_k, u_k) dla $k = 1, 2, \dots$, gdzie λ_k są wartościami własnymi oraz u_k są funkcjami własnymi. Wartości własne zależą od położenia i wielkości zniekształcenia \mathbf{D} . Dla rozważań prowadzonych w niniejszej pracy istotna jest relacja między wartościami własnymi a wielkością i lokalizacją zniekształcenia.

Głównym celem postawionym w pracy jest rozwiązanie zadania odwrotnego, które polega na lokalizowaniu zniekształcenia obszaru $\Omega_{\mathbf{D}}$ na podstawie wektora wartości własnych. Z uwagi na brak jednoznaczności rozwiązań zadania (4) zostało zdefiniowane nowe zadanie w innym obszarze, dla którego zagwarantowane jest istnienie przybliżonego rozwiązania zadania odwrotnego.

W celu wyznaczenia lokalizacji zniekształcenia definiujemy nowe odwzorowanie, którym jest warunkowa wartość oczekiwana położenia deformacji pod warunkiem, że znamy skończony ciąg wartości własnych będących rozwiązaniem zadania (4). Odwzorowanie to jest aproksymowane przez tak zwany ciąg aproksymujący. Ciągiem tym jest rodzina sieci neuronowych Elmana. Sieć jest budowana w sposób dynamiczny. Jej wielkość, rozumiana tutaj jako liczba neuronów kontekstowych oraz liczba neuronów w warstwie ukrytej, jest zależna od wielkości zbioru uczącego. Wraz z postępem procedury aproksymacji zwiększa się liczba elementów należących do zbioru uczącego, tym samym rośnie liczba neuronów w sieci neuronowej. Takie postępowanie pozwala wykazać, że stosowana metoda aproksymacji jest zbieżna.

Bibliografia

- [1] N. Baba, *A new approach for finding global minimum of error function of neural networks*, Neural Networks, 2, 1989, 367–374.
- [2] A. Barron, *Statistical Properties of Artificial Neural Networks*, Proceedings of the 28th IEEE Conference on Decision on Control, New York, 1989, 280–285.
- [3] M. Fisz, *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, Warszawa, 1967.
- [4] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin Heidelberg New York Tokyo, 1983.
- [5] M. Grzaneck, A. Laurain, K. Szulc, *Numerical algorithms for an inverse problem in shape optimization*, 6th International Conference on Inverse Problems in Engineering: Theory and Practice, Journal of Physics: Conference series, Vol. 135 (2008) 012047.
- [6] K. Hornik, M. Stinchcombe, H. White, *Multilayer feedforward networks are universal approximators*, Neural Networks, 2, 1989, 359–366.
- [7] K. Hornik, M. Stinchcombe, H. White, *Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks*, Neural Networks, 3, 1990.
- [8] L. Jackowska-Strumiłło, J. Sokołowski, A. Żochowski, A. Henrot, *On Numerical Solution of Shape Inverse Problem*, Computational Optimization and Application, 23, 2002, 231–255.
- [9] W. Kosiński, *Field singularities and wave analysis in continuum mechanics*, Mathematics and its Applications, PWN, Warszawa, 1986.
- [10] A. Laurain, K. Szulc, *Using self-adjoint extensions in shape optimization*, Proceedings of the 23rd IFIP TC7 Conference on System Modelling and Optimization, IFIP Advances in Information and Communication Technology, Vol. 312 (2009).
- [11] M. Stinchcombe, H. White, *Some measurability results for extrema of random functions over random sets*, University of California, San Diego 1992.
- [12] M. Stone, *Cross-validatory choice and assessment of statistical predictions*, Journal of the Royal Statistical Society, Series B, 36, 1974, 111–133.
- [13] H. White, J. Wooldridge, *Some results on sieve estimation with dependent observations*, Cambridge University Press, Now York, 1990.
- [14] H. White, *Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings*, University of California, San Diego 1990.
- [15] H. White, *Learning of artificial neural networks: A statistical perspective*, Neural Computation, 1, 1989, 425–464.
- [16] H. White, *Asymptotic theory for econometricians*, Orlando, 1984.
- [17] H. White, *A practical cross-validation method for multilayer feedforward networks*, Internal HNC Report, 1989.
- [18] M. Lipnicka, *Badania własności sieci neuronowych Elmana z uwzględnieniem szczególnych zastosowań*, Praca doktorska, Łódź, 2011.

Marta Lipnicka
Katedra Analizy Matematycznej i Teorii Sterowania
Wydział Matematyki i Informatyki
Uniwersytet Łódzki
E-mail: marta@math.uni.lodz.pl

Approximate localisation of imperfections in fixed domain

Abstract. In this paper we present a procedure for determining the approximate location of imperfection in a fixed domain. We define the spectral problem whose solutions are eigenvalues. These values depend on the location and the size of the imperfection. The main aim in this work is to find the solution of the inverse problem. It means that we find the location of the imperfection in our domain based on the vector of eigenvalues. For the inverse problem we don't have the uniqueness of the solutions so we Define a new problem in new domain.

For the new problem we obtain the existence of the approximate solution of the inverse problem. In order to determine the location of imperfection we define a new mapping. This mapping is defined as the conditional expectation of the location of imperfection, provided that we know the finite number of eigenvalues. The mapping is approximated by the Elman's neural networks. The networks are built in a dynamic way. Their size depends on the size of the learning set. The approximation method is convergent.

Keywords: Neural networks, Eigenvalues, Approximation, Conditional expectation.

(wpłynęło 27 lipca 2011 r.)