POLSKIE TOWARZYSTWO INFORMACJI PRZESTRZENNEJ
ROCZNIKI GEOMATYKI 2013 ○ TOM XI ○ ZESZYT 5(62)

ZASTOSOWANIE METODY OPTYMALIZACJI NIELINIOWEJ NELDERA-MEADA DO KONSTRUKCJI ODWZOROWAŃ KARTOGRAFICZNYCH O MOŻLIWIE NAJLEPSZYM ROZKŁADZIE ZNIEKSZTAŁCEŃ ODWZOROWAWCZYCH – NA PRZYKŁADZIE ODWZOROWANIA AZYMUTALNEGO

APPLICATION OF NELDER-MEAD NONLINEAR OPTIMIZATION METHOD FOR MINIMIZING MAP PROJECTION DISTORTION – DEMOSTRATED ON AN AZIMUTHAL PROJECTION

Kamil Jan Latuszek

Politechnika Warszawska, Wydział Geodezji i Kartografii, Zakład Kartografii

Słowa kluczowe: kartografia matematyczna, odwzorowania azymutalne, minimalizacja zniekształceń odwzorowawczych, kryterium Airy'ego, algorytm Neldera-Meada
Keywords: mathematical cartography, azimuthal projections, minimization of projection distortion, Airy's criterion, Nelder-Mead algorithm

Wstęp

Powierzchnia sfery jest nierozwijalna na płaszczyznę bez zniekształceń. W celu przedstawienia sferycznej powierzchni Ziemi na płaskiej powierzchni mapy stosowane są odwzorowania kartograficzne. Każde odwzorowanie kartograficzne niesie ze sobą zniekształcenia odwzorowawcze, jednakże w zależności od celu opracowania oraz odwzorowywanego obszaru, wybierane jest odwzorowanie o najbardziej odpowiednim rozkładzie zniekształceń.

Odwzorowanie takie może spełniać pewne kryteria podstawowe, jak wierne zobrazowanie kątów, pól powierzchni, czy wierne zachowywanie się odległości w zadanym kierunku. Może ono również nie spełniać żadnego z wymienionych kryteriów, w zamian spełniając określone kryteria integralne, całkowe – stanowiące pewien kompromis pomiędzy minimalizacją zniekształceń różnego rodzaju. Przykładem takiego kryterium całkowego jest podane w XIX wieku przez G.P. Airy'ego kryterium:

$$F = \sqrt{\frac{1}{2|S|}\iint\left[(m-1)^2+(n-1)^2\right]dS} = \min \tag{1}$$

W powyższym równaniu F oznacza miarę integralną zniekształceń odwzorowawczych, której minimum jest poszukiwane, m, n to ekstremalne skale długości w poszczególnych punktach, S denotuje obszar, dla którego kryterium ma być spełnione. Widać zatem, że odwzorowanie spełniające kryterium Airy'ego ma dla wybranego obszaru najmniejszą możliwą całkowitą sumę kwadratów odchyleń skal ekstremalnych od jedności.

Oprócz spełniania wybranych kryteriów względem charakteru i rozkładu zniekształceń odwzorowawczych, odwzorowanie kartograficzne spełnia z reguły pewne warunki geometryczne kształtu siatki kartograficznej, co na ogół wiąże się z przynależnością do pewnej klasy odwzorowań. Kryterium (1) można zatem rozważać dla odwzorowań azymutalnych, walcowych, stożkowych itd. Bez względu na ogólną charakterystykę kształtu siatki kartograficznej można również dążyć do uzyskania pewnych warunków geometrycznych dla wybranych linii albo grup linii parametrycznych. Przykładowo pożądany może być odpowiedni stosunek długości obrazu równika do długości obrazu południka osiowego, symetria siatki kartograficznej względem równika, odwzorowanie się bieguna północnego na pojedynczy punkt itd. Warunków tego typu można postulować więcej dla odwzorowań o bardziej dowolnym ogólnym kształcie siatek kartograficznych. Wprowadzane są one najczęściej poprzez odpowiednią modyfikację we wzorze ogólnym określającym dane odwzorowanie.

Odwzorowanie kartograficzne należące do danej klasy odwzorowań, spełniające wybrane warunki geometryczne wspomniane wyżej oraz spełniające pewne kryteria minimalizacji zniekształceń, jak np. (1) można uzyskiwać drogą analityczną lub przybliżonymi metodami numerycznymi. W przypadku metod przybliżonych mamy na ogół do czynienia z rozwiązywaniem odpowiedniego zadania optymalizacji nieliniowej (programowania nieliniowego), przy pomocy odpowiednio dobranego algorytmu numerycznego.

W przypadku optymalizacji rozkładu zniekształceń odwzorowawczych zadanie takie polega na znalezieniu minimum funkcji celu przyjętej jako kryterium oceny odwzorowania, np. kryterium (1), przez odpowiednie dobranie wartości parametrów opisujących współrzędne płaskie odwzorowania, pełniących rolę zmiennych decyzyjnych. Zbiór dopuszczalnych rozwiązań może być ograniczony. Zadanie optymalizacji odwzorowania kartograficznego można zapisać w następujący sposób:

$$F(\vec{a}) \longrightarrow \min,\ f_1(\vec{a}) < 0, f_2(\vec{a}) < 0, \ldots, f_n(\vec{a}) < 0 \tag{2}$$

W powyższym równaniu $F(\vec{a})$ jest minimalizowaną integralną miarą zniekształceń rozpatrywaną w danym kryterium, $\vec{a}$ jest wektorem parametrów opisujących odwzorowanie, $f_{i=1,\ldots,n}(\vec{a})$ są pewnymi funkcjami dla których określenie silnej nierówności względem zera ogranicza zbiór dopuszczalnych rozwiązań $\vec{a}$.

W praktyce funkcję celu $F(\vec{a})$ można zdefiniować jako skończoną sumę pewnych wartości stanowiącą przybliżenie rozpatrywanej integralnej miary zniekształceń odwzorowawczych. Ostateczna postać funkcji celu zależeć będzie od przyjętej klasy odwzorowań.

W niniejszym artykule rozpatrywane będzie rozwiązanie kryterium (1) bez ograniczeń nierównościowych dla odwzorowania azymutalnego normalnego sfery, z wykorzystaniem algorytmu Neldera-Meada.

Zastosowanie algorytmu Neldera-Meada do minimalizacji zniekształceń odwzorowawczych

Frank Canters (2002) optymalizował wybrane odwzorowania powierzchni całej sfery, przyjmując funkcję celu jako wartość zrewidowanej przez siebie miary Petersa:

$$F(\vec{a}) = \frac{1}{5000}\sum_{k=1}^{5000}\frac{|s'-s|}{|s'+s|} \tag{3}$$

W równaniu (3) s to odległość dwu losowo wybranych punktów na powierzchni sfery, s' to odległość tych punktów na powierzchni płaskiej odwzorowania. Punkty wybierane były losowo na powierzchni kontynentów (celem lepszego zobrazowania powierzchni kontynentów względem powierzchni oceanów). Jest to zatem skończona miara zniekształceń odwzorowawczych reprezentująca pewne przeciętne zniekształcenie tak zdefiniowanych odległości. Współrzędne optymalizowanego odwzorowania przyjęte były w postaci transformacji wielomianowej współrzędnych wyjściowych tego odwzorowania:

$$X = \sum_{i=0}^{n}\sum_{j=0}^{n-1} a_{i,j}x^i y^j \tag{4}$$

$$Y = \sum_{i=0}^{n}\sum_{j=0}^{n-1} a'_{i,j}x^i y^j \tag{5}$$

W równaniach (4-5) X, Y oznaczają współrzędne po transformacji, x, y są współrzędnymi odwzorowania macierzystego, współczynniki a_{ij}, a'_{ij} są współczynnikami dwuparametrowej transformacji wielomianowej.

Optymalizowanymi, ze względu na wartość funkcji celu (3) zmiennymi decyzyjnymi były współczynniki wielomianów (4-5). Wstępnie na wartości współczynników powyższej transformacji nakładane były odpowiednie warunki pozwalające uzyskać wymagane własności geometryczne siatki kartograficznej. Przykładowo, dla zachowania symetrycznego rozłożenia punktów względem osi Ox siatki macierzystej, nieparzyste potęgi y w (4) oraz parzyste potęgi y w (5) muszą zostać usunięte. Analogicznie dla uzyskania symetrii względem osi Ox parzyste potęgi x w (4) oraz nieparzyste potęgi x w (5) muszą zniknąć. Nakładanie tego typu ograniczeń geometrycznych jest ważne dla prawidłowego odwzorowywania całego globu (Canters, 2002).

Zadanie optymalizacji odwzorowania opisanego przez (4-5) dla funkcji celu (3) sprowadza się do znalezienia takich wartości współczynników a_{ij}, a'_{ij}, dla których po podstawieniu (4-5) do (3) wartość prawej strony (3) będzie minimalna, przy założonych ograniczeniach na wartości tych współczynników. Do rozwiązania tego zadania Canters posłużył się algorytmem Neldera-Meada (ang. *downhill simplex method*) optymalizacji nieliniowej, nazywanej czasem w polskiej literaturze algorytmem pełzającego sympleksu. Algorytm ten został przedstawiony i opisany w pracy (Nelder, Mead, 1965). Nie należy mylić tego algorytmu z algorytmem sympleks programowania liniowego.

Sympleks to figura matematyczna stanowiąca n+1 wierzchołków w n-wymiarowej przestrzeni. Każde n wierzchołków w sympleksie jest liniowo niezależne. Dla n optymalizowanych zmiennych decyzyjnych generowany jest początkowy sympleks składający się z n+1 wierzchołków, mających n współrzędnych. Współrzędne wierzchołka odpowiadają wartościom zmiennych decyzyjnych w tym wierzchołku. W niniejszym artykule sympleks będzie oznaczony jako macierz $A_{n+1,n}$

$$A_{n+1,n} = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n+1,1} & \cdots & a_{n+1,n} \end{bmatrix} \tag{6}$$

Pojedynczy wierzchołek sympleksu jest jednym z wierszy macierzy A, to znaczy wektorem:

$$\vec{a_i} = \left[a_{i,1}, a_{i,2}, \ldots, a_{i,n}\right] \tag{7}$$

Dla każdego takiego wierzchołka można obliczyć wartość funkcji celu zależnej od n zmiennych decyzyjnych $F(\vec{a_i})$. Oznaczmy przez $\vec{a_L}$ – wierzchołek sympleksu o najmniejszej wartości funkcji celu, – wierzchołek o największej wartości funkcji celu oraz $\vec{a_{h-1}}$ wierzchołek o drugiej największej wartości funkcji celu. Wartości indeksów L, h, h–1 wskazują odpowiedni wiersz w (6).

Istotą metody jest przesuwanie w każdej iteracji wierzchołka „najgorszego” (o największej wartości funkcji celu, $\vec{a_h}$) na taką nową pozycję w n-wymiarowej przestrzeni, aby jego wartość się poprawiła. Wierzchołek przesuwany jest w kierunku środka ciężkości (centroidu) $\vec{a_c}$ pozostałych wierzchołków. Odległość przesunięcia zależy od stopnia poprawy wartości funkcji celu dla przemieszczonego wierzchołka.

Wstępnym etapem każdej iteracji w algorytmie jest znalezienie wierzchołków $\vec{a_L}, \vec{a_h}, \vec{a_{h-1}}$ oraz wyznaczenie środka ciężkości $\vec{a_c}$ dla wszystkich wierzchołków za wyjątkiem $\vec{a_h}$:

$$\vec{a_c} = \frac{1}{n}\left[\left(\sum_{i=1}^{n+1}\vec{a_i}\right) - \vec{a_h}\right] \tag{8}$$

Następnie wykonywane jest „odbicie” (ang. *reflection*) – polegające na przesunięciu prostoliniowym wierzchołka w kierunku $\vec{a_c}$ tak, aby znalazł się po jego drugiej stronie:

$$\vec{r} = \vec{a_c} + \alpha(\vec{a_c} - \vec{a_h}), \quad 0 < \alpha \tag{9}$$

Dla nowej pozycji „odbitego” względem centroidu wierzchołka $\vec{r}$ obliczona zostaje wartość funkcji celu $F(\vec{r})$. Ta wartość porównywana jest z wartościami funkcji celu w wierzchołkach $\vec{a_L}, \vec{a_h}, \vec{a_{h-1}}$. Jeśli nowa wartość funkcji celu jest mniejsza niż w wierzchołku $\vec{a_L}$, wykonywane jest tzw. „rozciągnięcie” (ang. *expansion*). Przyjęta jest wówczas heurystyka, że kierunek przesunięcia jest obiecujący ze względu na zmniejszenie wartości funkcji celu i należy wykonać dalsze przesunięcie w tym kierunku:

$$\vec{e}=\vec{a_c}+\gamma(\vec{a_c}-\vec{a_h}),\ \gamma>1, \gamma>\alpha \tag{10}$$

Wierzchołek $\vec{a_{i=h}}$ jest wówczas wymieniony w sympleksie (6) na $\vec{e}$, jeśli wartość *F(e)* jest mniejsza od *F(a_L)*, lub na $\vec{r}$ w przeciwnym wypadku. Jeśli wartość funkcji celu w $\vec{r}$ jest większa lub równa wartości funkcji celu w $\vec{a_L}$ ale mniejsza niż w $\vec{a_{h-1}}$, wówczas wierzchołek $\vec{a_h}$ wymieniany jest również na $\vec{r}$. Natomiast jesli funkcja celu w wierzchołku $\vec{r}$ osiąga większą lub równą wartość, w stosunku do $\vec{a_{h-1}}$, ale mniejszą niż w $\vec{a_h}$, wówczas dokonywane jest tzw. „ściągnięcie" (ang. *contraction*) oraz wierzchołek $\vec{a_h}$ wymieniany jest na $\vec{c}$:

$$\vec{c}=\vec{a_c}+\beta(\vec{r}-\vec{a_c}),\ \ 0<\beta<1, \beta<\alpha \tag{11.1}$$

Jeśli natomiast $\vec{r}$ okazuje się „nielepszy" w sensie minimalizacji funkcji celu od $\vec{a_h}$, a zatem operacja odbicia nie poprawiła wartości funkcji celu w tym wierzchołku, wówczas również wykonane jest ściągnięcie, jednakże po stronie wierzchołka $\vec{a_h}$:

$$\vec{c}=\vec{a_c}+\beta(\vec{a_h}-\vec{a_c}),\ \ 0<\beta<1,\ \ \beta<\alpha \tag{11.2}$$

po czym jeśli wierzchołek $\vec{c}$ jest lepszy niż $\vec{a_h}$ – następuje wymiana $\vec{a_h}$ na $\vec{a_c}$, w przeciwnym wypadku wszystkie inne wierzchołki są uśredniane z wierzchołkiem $\vec{a_L}$, a $\vec{a_h}$ wymieniony zostaje na $\vec{c}$:

$$\bigwedge_{i\in\{1,\dots,n+1\}} \vec{a_i}:=(\vec{a_i}+\vec{a_L})/2 \tag{12}$$

Nelder i Mead eksperymentalnie wywnioskowali, że na ogół najszybszą zbieżność uzyskuje się dla $\alpha = 1$, $\beta = 0{,}5$, $\gamma = 2$. W początkowym stadium działania algorytmu, jeśli sympleks jest daleki od minimum – dzięki operacji rozciągnięcia zwiększa swój rozmiar, co implikuje także zwiększanie odległości kolejnych odbić i rozciągnięć w kolejnych iteracjach, co przyspiesza proces zbliżania się figury sympleksu w okolice minimum. Gdy sympleks położony jest blisko minimum, jego rozmiar zaczyna się stopniowo zmniejszać dzięki operacjom kontrakcji i uśredniania pozycji wierzchołków z wierzchołkiem najlepszym $\vec{a_L}$. W efekcie sympleks „zapada" się do ostatecznego minimum. Za znalezione wartości zmiennych decyzyjnych można ostatecznie przyjąć współrzędne wierzchołka $\vec{a_L}$ po zakończonej ostatniej iteracji. Kryterium zakończenia procesu iteracyjnego może być np. różnica $F(\vec{a_L})-F(\vec{a_h})$

Optymalizacja odwzorowania azymutalnego celem spełnienia kryterium Airy`ego

W pracy (Gdowski, 1967) Gdowski podaje rozwiązanie kryterium Airy'ego (1) dla odwzorowania azymutalnego w położeniu normalnym, uzyskane drogą analityczną, z wykorzystaniem rachunku wariacyjnego. Kryterium (1) można dla odwzorowania azymutalnego normalnego sfery jednostkowej zapisać w postaci:

$$\delta=\int_{\alpha}^{\beta}\left[\left(\frac{dr}{d\theta}-1\right)^2+\left(\frac{r}{\sin\theta}-1\right)^2\right]\sin\theta d\theta \quad \text{gdzie: } m=\frac{dr}{d\theta} \quad n=\frac{r}{\sin\theta} \tag{13}$$

przy czym $r = r(\theta)$ to promień obrazu równoleżnika, m, n są ekstremalnymi skalami długości w poszczególnych punktach, granice całkowania $\theta_{min} = \alpha$, $\theta_{max} = \beta$ wskazują granice obszaru optymalizowanego (jego położenie), $\theta = \pi/2 - \varphi$ jest odległością sferyczną od bieguna północnego. Uzyskane drogą analityczną odwzorowanie spełniające (13) ma postać:

$$x = r(\theta)\cos(\lambda) \qquad y = r(\theta)\sin(\lambda) \tag{14}$$

$$r(\theta) = A tg\frac{\theta}{2} + B ctg\frac{\theta}{2} + 2 ctg\frac{\theta}{2}\lg\sec\frac{\theta}{2}$$

$$A = \frac{2\left(\lg\sec\frac{\beta}{2} - \lg\sec\frac{\alpha}{2}\right)}{tg^2\frac{\beta}{2} - tg^2\frac{\alpha}{2}} \qquad B = \frac{2\left(tg^2\frac{\alpha}{2}\lg\sec\frac{\beta}{2} - tg^2\frac{\beta}{2}\lg\sec\frac{\alpha}{2}\right)}{tg^2\frac{\beta}{2} - tg^2\frac{\alpha}{2}}$$

Aby uzyskać odwzorowanie azymutalne spełniające kryterium Airye'go metodami numerycznymi, należy odpowiednio sformułować zadanie programowania nieliniowego, to jest wprowadzić odpowiadającą kryterium (13) funkcję celu. Rozpatrzymy dwie miary, odniesione do odwzorowania sfery jednostkowej, z których pierwsza jest postaci:

$$F(\vec{a}) = \sum_{k=1}^{n_k}\left[\left(\frac{s'_p(\theta_k) - s_p}{s_p}\right)^2 + \left(\frac{s'_r(\theta_k) - s_r}{s_r}\right)^2\right]\sin(\theta_k) \tag{15}$$

$$\theta_k = a + k\Delta - \Delta/2 \qquad s_p = s_r = \Delta, \qquad \Delta = (\beta - \alpha)/n_k$$

$$s'_r(\theta_k) = r(\theta_k)\Delta/\sin(\theta_k) \qquad s'_p(\theta_k) = \left|r(\theta_k - \Delta/2) - r(\theta_k + \Delta/2)\right|$$

W (15) n_k to liczba odcinków południków, równa liczbie odcinków równoleżników, dla których liczone jest względne zniekształcenie długości, θ_k jest odległością sferyczną od bieguna północnego, wyznaczoną dla punktu środkowego kolejnych odcinków południków oraz dla punktów znajdujących się na odpowiadających im odcinkach równoleżników, Δ jest długością fragmentów równoleżników i południków $s_r = s_p$ na powierzchni sfery, α, β to granice obszaru optymalizowanego podane w (13), $r(\theta)$ jest promieniem wodzącym równoleżników w odwzorowaniu azymutalnym, s'_r, s_p są długościami fragmentów równoleżników i południków na powierzchni odwzorowania, $\vec{a}$ jest wektorem parametrów opisujących odwzorowanie.

Funkcja (15) w założeniu odpowiada zrewidowanej mierze Petersa dla odwzorowania azymutalnego mającego spełnić kryterium Airy'ego. Ponieważ w kryterium Airy'ego interesuje nas minimalizacja zniekształceń w kierunkach skal ekstremalnych, a w odwzorowaniu azymutalnym są to kierunki skal parametrycznych, stąd wniosek, że należy minimalizować zniekształcenia długości fragmentów południków i równoleżników. Uwzględniając południkowy rozkład zniekształceń odwzorowawczych, wnioskujemy, że fragmenty południków między takimi samymi szerokościami geograficznymi, zniekształcają się tak samo, a ich długość na powierzchni obrazu odpowiada różnicy długości promieni wodzących równoleżników, między którymi te odcinki południków się znajdują. Fragmenty południków powinny mieć łączną długość południka między granicami α, β i posiadać jedynie wspólne końce. Fragmenty równoleżników zaś powinny przechodzić przez środki fragmentów południków, a ich długość na powierzchni obrazu wyraża się jako długość fragmentu łuku obrazu równoleżnika. W ten sposób uzyskamy równomierne rozmieszczenie odcinków linii parametrycznych. Długości fragmentów równoleżników i południków powinny być identyczne na powierzchni oryginału, stąd stała wartość $\Delta = s_p = s_r$. Zgodnie z zapisem (13) zniekształcenia długości odcinków równoleżników i południków należy przemnożyć współczynnikiem ważącym sin (θ_k), gdzie θ_k to średnia szerokość geograficzna fragmentu południka, a zarazem fragmentu równoleżnika przechodzącego przez niego prostopadle. Nie wprowadza się czynnika losowego w doborze odcinków południków i równoleżników, gdyż w odwzorowaniu azymutalnym zniekształcenia zmieniają się stopniowo w kierunku południkowym. Należy natomiast dobrać odpowiednio mały interwał Δ, gdyż dla dłuższych odległości na powierzchni obrazu w każdym odwzorowaniu zachodzić może kompensacja zniekształceń ujemnych i dodatnich. Za przyjęciem takiej funkcji celu przemawia związek zniekształceń długości fragmentów południków i równoleżników ze zniekształceniami w kierunku skal ekstremalnych oraz fakt, że pomimo skończonej liczby sumowanych elementów n_k, zniekształcenia wszystkich punktów w zadanym pasie wpływają na obliczoną wartość funkcji celu – co dobrze odpowiada integralnemu charakterowi całki (13).

Alternatywną funkcją celu może być całka w (13), policzona dla skończonej liczby punktów, równomiernie rozłożonych wzdłuż jednego południka:

$$F(\vec{a}) = \sum_{k=0}^{n_k} \left[\left(r'(\theta_k) - 1\right)^2 + \left(\frac{r(\theta_k)}{\sin(\theta_k)} - 1 \right)^2 \right] \sin(\theta_k) \tag{16}$$

$$\theta_k = \alpha + k\Delta \qquad \Delta = (\beta - \alpha)/n_k$$

Symbole użyte do opisu powyższych zależności należy rozumieć tak, jak dla poprzednio sformułowanej funkcji celu, z tą różnicą, że Δ odpowiada odstępom między punktami, dla których wyznaczane są ekstremalne skale długości, zamiast środkom fragmentów równoleżników i południków, $r'(\theta)$ oznacza pochodną promienia $r(\theta)$ względem θ. Taka funkcja celu odpowiada całkowaniu numerycznemu (13), bez mnożenia wartości całki przez przedział $d\theta$, gdyż przy równym podziale Δ, minimalizacja (16) równoważy minimalizację całki numerycznej (13). Za zastosowaniem takiej funkcji celu przemawia bezpośredni związek pomiędzy wartością całki i jej przybliżeniem numerycznym.

Współrzędne płaskie w optymalizowanym numerycznie odwzorowaniu azymutalnym można przedstawić w postaci:

$$x = r(\theta)\cos(\lambda)$$

$$y = r(\theta)\sin(\lambda) \tag{17}$$

$$r(\theta) = a_0 r_0(\theta) + \sum_{i=1}^{n} a_i \theta^{i-1}$$

gdzie promień $r(\theta)$ dany jest w postaci kombinacji liniowej pewnego promienia odwzorowania azymutalnego wziętego jako startowe $r_0(\theta)$ i kolejnych potęg θ. Współczynniki a_0, $a_1,\ldots, a_n$ tej kombinacji liniowej są zmiennymi decyzyjnymi w zadaniu optymalizacji odwzorowania kartograficznego oraz stanowią współrzędne wektora $\vec{a}$ w równaniach (15-16). Dla funkcji celu opisanej w (15) mamy po podstawieniu współrzędnych (17):

$$F(\vec{a}) = \sum_{k=1}^{n_k} \left\{ \frac{\sin(\theta_k)}{\Delta^2} \left(a_0 \left| r_0\left(\theta_k + \frac{\Delta}{2}\right) - r_0\left(\theta_k - \frac{\Delta}{2}\right) \right| + \sum_{i=1}^{n} a_i \left[\left(\theta_k + \frac{\Delta}{2}\right)^{i-1} - \left(\theta_k - \frac{\Delta}{2}\right)^{i-1} \right] - \Delta \right)^2 \right\}$$
$$+ \sum_{k=1}^{n_k} \left\{ \sin(\theta_k) \left[\frac{1}{\sin(\theta_k)} \left(a_0 r_0(\theta_k) + \sum_{i=1}^{n} a_i \theta_k^{\,i-1} \right) - 1 \right]^2 \right\} \tag{18}$$

Pochodna promienia $r'(\theta)$ w (17) przedstawia się w postaci:

$$r'(\theta) = a_0 r'_0(\theta) + \sum_{i=2}^{n} (i-1) a_i \theta^{i-2} \tag{19}$$

Dla funkcji celu (16) po podstawieniu (17) i (19) uzyskamy:

$$F(\vec{a}) = \sum_{k=1}^{n_k} \left\{ \left(a_0 r'_0(\theta_k) + \sum_{i=2}^{n} (i-1) a_i \theta_k^{\,i-2} - 1 \right)^2 + \left[\frac{1}{\sin(\theta_k)} \left(a_0 r_0(\theta_k) + \sum_{i=1}^{n} a_i \theta_k^{\,i-1} \right) - 1 \right]^2 \right\} \sin\theta_k \tag{20}$$

Zarówno w (18) jak i w (20) funkcja celu jest wielomianem postaci:

$$F(\vec{a}) = \sum_{k=1}^{n_k} \left\{ \left(a_0 u_{k,0} + \sum_{i=1}^{n} a_i u_{k,i} - 1 \right)^2 + \left(a_0 u'_{k,0} + \sum_{i=1}^{n} a_i u'_{k,i} - 1 \right)^2 \right\} f_k = \sum_{k=1}^{2n_k} \left(\sum_{i=0}^{n} a_i v_{k,i} - l_k \right)^2 \tag{21}$$

Przy czym współczynniki $u_{k,i}$, $u`_{k,i}$, f_k, $v_{k,i}$, l_k są stałe, a zmienne decyzyjne są współczynnikami a_i – współrzędnymi wektora $\vec{a}$.

Na potrzeby niniejszego opracowania optymalizowano obydwie funkcje celu (18) oraz (20) dla zmiennych decyzyjnych będących współczynnikami kombinacji liniowej (17). Wykorzystano algorytm Neldera-Meada do znalezienia minimum tych funkcji. Za kryterium zakończenia obliczeń przyjęto liczbę iteracji. Odwzorowaniem początkowym $r_0(\theta)$ było odwzorowanie równopolowe:

$$r_0(\theta) = 2\sin(\theta/2) \tag{22}$$

Sympleks początkowy przyjęto jako:

$$A_{n+1,n} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0.05 & 0 & \cdots & 0 \\ 1 & 0 & 0.05 & \cdots & 0 \\ \vdots & 0 & 0 & \ddots & 0 \\ 1 & 0 & 0 & \cdots & 0.05 \\ 0.95 & 0 & 0 & \cdots & 0 \end{bmatrix}_{n+1,n} \tag{23}$$

Zatem pierwszy wierzchołek sympleksu odpowiada odwzorowaniu (22), czwarty odwzorowaniu $r_0(\theta) = 2 \sin(\theta/2) + 0{,}05\ \theta^2$ itp.

W tabeli zaprezentowano wyniki działania algorytmu dla różnej liczby użytych współczynników rozwinięcia (17), różnej liczby wydzielonych w funkcji celu (18) fragmentów południków i równoleżników n_k lub punktów pomiaru odchyleń skal od jedności w (20), podając przy tym liczbę wykonanych iteracji, oraz wartość funkcji celu (wyniki dotyczą obszaru zawartego między $\theta_1 = a = 35°$, $\theta_2 = \beta = 41°$, co odpowiada pasowi równoleżnikowemu, w którym leży Polska). Wartości funkcji celu znormalizowano dzieląc przez liczbę n_k – co oznacza, że podane wartości wyrażają średnią sumę kwadratów odchyleń skal ekstremalnych od jedności w pojedynczym punkcie. Liczba wykonanych iteracji jest orientacyjną wartością, dla której przy podanym sympleksie startowym (23) wynik przestaje ulegać istotnej zmianie. W ostatniej kolumnie podano wyniki dla funkcji celu $F'(a)$ wziętej jak w (20), dla $n_k' = 1000$ punktów na południku. Wcześniej ustalono, że dla odwzorowania analitycznego i dla takiej liczby punktów – dokładność całkowania numerycznego jest odpowiednia, by reprezentować rzeczywistą całkę z dokładnością do 5 cyfr znaczących. Obliczenia prowadzono z dokładnością 4 miejsc po przecinku dla optymalizowanych współczynników rozwinięcia (17). Prowadzenie obliczeń z taką dokładnością pozwala na osiągnięcie minimum badanego kryterium z dokładnością około 1/6000, co odpowiada zmianie zniekształceń na poziomie pojedynczych milimetrów na kilometr, co jest zaniedbywalne, biorąc pod uwagę wielkość zniekształceń dla odwzorowania analitycznego (14), dochodzącą do metrów na kilometr wzdłuż równoleżników. W pierwszej kolumnie tabeli dla odwzorowania analitycznego wpisano literę „A”, natomiast różne odwzorowania zanumerowano od 1 do 16.

Analizując zawarte w tabeli wyniki można stwierdzić, że już dla odwzorowania o 3 optymalizowanych współczynnikach (numery 1-2), optymalizowanego względem funkcji celu uwzględniającej podział południka na 10 części, funkcja celu może osiągnąć wartość na poziomie poniżej 1% różnicy w stosunku do odwzorowania uzyskanego drogą ścisłą. Nieznaczny wzrost wartości całki $F(\alpha)$ dla wzrastającej częstości podziału n_k spowodowany jest większym udziałem szerokości geograficznych na skraju obszaru w tworzeniu wartości funkcji celu – która uwzględnia duże wartości zniekształceń dla tych szerokości po podniesieniu do kwadratu. Dla funkcji celu (18) nie obserwowany jest przy tym spadek tych wartości przez kompensację zniekształceń ujemnych i dodatnich, gdyż częstość podziału jest wystarczająco duża dla $n_k = 10$, ponadto w uzyskiwanych odwzorowaniach przeważający wpływ na zniekształcenia miała skala równoleżnikowa – o stałej wartości bez względu na długość fragmentu równoleżnika. Zastosowanie funkcji celu podanej jak w (18) jest równoważne (20) ze względu na braną pod uwagę dokładność. Obydwie funkcje dają poprawne wyniki, porównywalne z rozwiązaniem analitycznym. Dla obszaru o wskazanym rozmiarze i położeniu okazuje się wystarczające wykorzystanie 4 początkowych współczynników w kombinacji liniowej (17), (rozwiązania 3-8 w tabeli) przy podziale $n_k = 30$ daje to możliwość szybkiego wyznaczenia odwzorowania drogą numeryczną z dokładnością około 0,02% wartości funkcji celu.

Tabela. Odwzorowania azymutalne uzyskane drogą optymalizacji nieliniowej w porównaniu z rozwiązaniem analitycznym „A”

i p.	a_0	a_1	a_2	a_3	a_4	a_5	a_6		i iczba iteracji			$F(x)/1000$
										Która funkcja celu?		$nk' = 1000$
	$(*r_1(\theta))$	$(*1)$	$(*\theta)$	$(*\theta^2)$	$(*\theta^3)$	$(*\theta^4)$	$(*\theta^5)$	n_k		↓	$F(x) / n^k$	
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	10	Nd.	18	6,6031E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	30	Nd.	18	6,6630E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	100	Nd.	18	6,6698E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	1000	Nd.	18	6,6705E-05	6,6705E-05
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	10000	Nd.	18	6,6705E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	10	Nd.	20	6,6033E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	30	Nd.	20	6,6630E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	100	Nd.	20	6,6698E-05	Nd.
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	1000	Nd.	20	6,6705E-05	6,6705E-05
A	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	Nd.	10000	Nd.	20	6,6705E-05	Nd.
1	-0,0047	0,0471	1,0039	Nd.	Nd.	Nd.	Nd.	10	300	18	6,618E-05	6,685E-05
2	-0,0047	0,0471	1,0039	Nd.	Nd.	Nd.	Nd.	10	300	20	6,618E-05	6,685E-05
3	2,5424	0,0769	3,6773	0,2067	Nd.	Nd.	Nd.	10	500	18	6,603E-05	6,671E-05
4	-2,4808	0,0762	3,6125	0,2017	Nd.	Nd.	Nd.	10	500	20	6,604E-05	6,671E-05
5	-2,4025	0,0752	3,5299	0,1954	Nd.	Nd.	Nd.	20	500	18	6,655E-05	6,672E-05
6	-2,3875	0,0751	3,5142	0,1941	Nd.	Nd.	Nd.	20	500	20	6,654E-05	6,671E-05
7	-2,3775	-0,075	3,5036	0,1933	Nd.	Nd.	Nd.	30	500	18	6,663E-05	6,671E-05
8	-2,3709	0,0749	3,4966	0,1928	Nd.	Nd.	Nd.	30	500	20	6,665E-05	6,671E-05
9	1,0657	0,0759	0,0639	0,1959	0,1421	Nd.	Nd.	10	700	18	6,604E-05	6,671E-05
10	1,0528	0,0752	0,0736	0,1911	0,1392	Nd.	Nd.	10	700	20	6,604E-05	6,671E-05
11	1,0487	0,0743	0,0737	-0,185	0,1359	Nd.	Nd.	20	700	18	6,654E-05	6,671E-05
12	1,0487	0,0742	0,0728	0,1838	0,1354	Nd.	Nd.	20	700	20	6,654E-05	6,671E-05
13	1,0499	0,0741	0,0711	-0,183	0,135	Nd.	Nd.	30	700	18	6,664E-05	6,671E-05
14	1,0498	-0,074	0.0709	0,1825	0,1347	Nd.	Nd.	30	700	20	6,665E-05	6,671E-05
15	-2,3578	0,0747	3,4828	0,1917	Nd.	Nd.	Nd.	1000	700	18	6,671E-05	6,6705E-05
16	0,9921	0,0557	0,0215	0,0461	0,0696	0,0446	0,0189	30	1000	20	6,664E-05	6,671E-05

Podsumowanie

Współcześnie możliwości obliczeniowe komputerów pozwalają na konstruowanie oraz optymalizację odwzorowań różnego typu, spełniających różne kryteria minimalizacji zniekształceń kartograficznych. Uzyskane drogą numeryczną odwzorowania są rozwiązaniami przybliżonymi. Ich przewaga nad rozwiązaniami ścisłymi polega na tym, że poprawnie skonstruowany algorytm optymalizacji odwzorowania kartograficznego dla danego kryterium, może zostać zastosowany lub łatwo zmodyfikowany celem zastosowania dla innej, wybranej klasy odwzorowań oraz na potrzeby innego kryterium. Zmiana kryterium wiąże się z redefinicją funkcji celu, zmiana klasy odwzorowania wiąże się natomiast z modyfikacją formuł odwzorowawczych na współrzędne płaskie i tym samym ustaleniu zmiennych decyzyjnych dla funkcji celu. W omawianym w artykule przykładzie odwzorowania azymutalnego spełniającego kryterium Airy'ego, zastosowano z powodzeniem algorytm optymalizacji nieliniowej Neldera-Meada. Obydwie zaproponowane funkcje celu (15, 16) dla n_k dążącego do nieskończoności dążą do dokładnej wartości funkcji podanej w kryterium (13), jednakże dla uzyskania dokładności rzędu 0,02% spełnienia kryterium wystarczy przyjąć n_k = 30 oraz wziąć pod uwagę cztery pierwsze wyrazy rozwinięcia (17).

Literatura

Canters F., 2002: Small-scale Map Projection Design. Londyn, Nowy York: Taylor & Francis.

Gdowski B., 1967: Kryterium Airy i Fioriniego oraz ich uogólnienia w zastosowaniu do klasy normalnych odwzorowań azymutalnych, *Geodezja i Kartografia* rocznik XVI nr 4.

Nelder J.A., Mead R., 1965: A simplex method for function minimization, *Computer Journal* 7: 308-313.

Abstract

The search for map projections with least possible distortion, satisfying selected criteria which integrate different measures of distortion, is one of the more important tasks of cartography. In the nineteenth century, many integral based criterions have been proposed, minimization of which is considered as achieving an optimal distortion pattern for a given projection. In the present time of mass computerization and constantly rising computation speed, popularity of numerical solutions of the mentioned criteria has risen. These numerical solutions are achieved by application of nonlinear optimization methods.

A nonlinear function minimization method proposed by Nelder and Mead (Nelder and Mead, 1965) was used to optimize map projections of the spherical globe for small scale mapping by Frank Canters (2002). Canters optimized projections of the whole globe, for which flat coordinates were given by fifth order polynomials. Parameters of these polynomials were either longitude and latitude on the globe or flat coordinates of a given parent projection. The objective function was the revised Peters measure of distortion (Canters, 2002), which is a finite distortion measure comparing distance between two given points on the globe with their distance on the map, for a large set of randomly chosen points.

In the present study, Nelder-Mead algorithm is used to minimize distortion of an azimuthal projection of the sphere in the normal aspect, so that it will satisfy Airy's criterion. The obtained solution will be then compared with an analytical-strict solution for this criterion, as given by Gdowski (1967). The parallel radius in the formulas describing flat coordinates of the optimized projection is written as a linear combination of the parent projections radius and a power series of θ, which denotes spherical distance from the north pole. Optimized variables will be the coefficients of the said linear combination with flat coordinates of the optimized projection given as:

$$x = r(\theta)\cos(\lambda)$$

$$y = r(\theta)\sin(\lambda)$$

$$r(\theta) = a_0 r_0(\theta) + \sum_{i=1}^{n} a_i \theta^{i-1}$$

As the objective function, we consider the value of adequately modified Peters measure, so that its minimization corresponds with the minimization of the sum of the squares of the errors of scale factors in principal directions, for the given region, as stated in the Airy's criterion.

The author plans to continue his research of nonlinear optimization methods for other projections and distortion minimization criterions.

mgr inż. Kamil Jan Latuszek
kamil.latuszek@wp.pl