Seria zadań domowych - arytmetyka fl

Proszę rozwiązać w formie pisemnej zadania odpowiednio oznaczone (słowo pisemnie) czyli zad 1 i 2 na piątek **18 marca 2011** - sprawdzone i ocenione mogą być tylko niektóre z tak oznaczonych i tylko te będą punktowane.

Zadanie 1 **(pisemnie)** Chcemy obliczyć funkcję $f(x) = \exp(10^7 x)$ w arytmetyce pojedynczej precyzji. Policz (przybliżony i względny) współczynnik uwarunkowania zadania dla $x \in [-10, 10]$ i określ czy obliczanie w tej arytmetyce fl wartości f dla $x = -6$ jest dobrze uwarunkowane ze względu na błąd względny?

Zadanie 2 **(pisemnie)** Chcemy w fl obliczyć $f(x) = 2x - \sqrt{4x^2 + 7}$. Rozpatrzmy dwa algorytmy: w pierwszym liczymy:

$f := \sqrt{4x^2 + 7}; f := 2x - f;$

a w drugim:

$f := \sqrt{4x^2 + 7}; f := -7.0/(2x + f);$

Który z nich należy zastosować do obliczenia w arytmetyce fl pojedynczej precyzji $f(10^7)$? (Podać krótkie uzasadnienie - 1-2 zdania - nie trzeba formalnie dowodzić oszacowań)

Zadanie 3 Pokaż, że następujący naturalny algorytm obliczania: iloczynu skalarnego dwóch wektorów $x, y \in \mathbb{R}^M$ tzn. $x^T y = \sum_{k=1}^{M} x(k)y(k)$ jest numerycznie poprawny tzn. obliczamy

```
s=x(1)*y(1);
for k=2:M,
  s=s+x(k)*y(k);
endfor
```

Zadanie 4 Pokaż, że następujący naturalny algorytm obliczania iloczynu macierz kwadratowej $A \in \mathbb{R}^{M,M}$ i wektora $x, \in \mathbb{R}^M$ tzn. $y = Ax$: jest numerycznie poprawny:

```
for j=1:M,
 y(j)=0;
 for k=1:M,
   y(j)=y(j)+A(j,k)*x(k);
 endfor
endfor
```

Wskazówka: $w(j)$ jest iloczynem skalarnym $j-$tego wiersza A z x.

Zadanie 5 (trudne) Pokaż, że następujący naturalny algorytm obliczania: cosinusa kąta dwóch wektorów $x, y \in \mathbb{R}^M$ tzn. $\cos(x, y) = \frac{x^T y}{\sqrt{x^T x}\sqrt{y^T y}}$ jest numerycznie poprawny:

(a) $a = \sqrt{\sum_{k=1}^{M} x_k^2}$, $b = \sqrt{\sum_{k=1}^{M} y_k^2}$

(b) $c = \sum_{k=1}^{M} x_k y_k$

(c) $w = c/(a * b)$

Zakładamy że wynikiem funkcji pierwiastek w arytmetyce fl jest $fl(\sqrt{x}) = x(1+\epsilon)$ dla $|\epsilon| \leq \nu$ i obliczanie iloczynu skalarnego $w^T z$ jest standardowe jak w zadaniu na NP obliczania iloczynu skalarnego.

Szkice rozwiązań niektórych zadań

Zad. 1 Uwarunkowanie - wzór z ćwiczeń. W arytmetyce pojedynczej precyzji uwarunkowanie za duże.

Zad. 2 Algorytm drugi lepszy - dla dużych dodatnich xw algorytmie pierwszym odejmujemy bliskie sobie wartości - i ryzykujemy dużą utratę dokładności. W drugim algorytmie nie ma tego ryzyka.

Zad. 3 Trzeba pokazać, że dla $s = f(x,y) = x^T y$ ten algorytm w fl zwróci takie $\tilde{s}$, że istnieją $\hat{x}, \hat{y}$ takie, że $\hat{s} = f(\hat{x}, \hat{y})$ takie, że

$$\frac{|s - \hat{s}|}{|\hat{s}|} \leq K_2 \nu + O(\nu^2)$$

z $\frac{\sum_k (|x_k - \hat{x}_k| + |y_k - \hat{y}_k|)}{\sum_k (|x_k| + |y_k|)} \leq K_1 \nu + O(\nu^2)$. Normę pierwszą można zastąpić inną. Stałe K_1, K_2 nie powinny zależeć od x, y i nie być *za duże* względem ν. (Oczywiście to trochę nieprecyzyjne).

Otrzymujemy po wykonaniu działań w fl (δ_k pojawia się w wyniki mnożeń w fl, ϵ_j w wyniku dodawania w fl), że

$$\begin{aligned} fl(s) &= (\ldots(x_1 y_1 (1+\delta_1) + x_2 y_2 (1+\delta_2))(1+\epsilon_2) + x_3 y_3 (1+\delta_3))(1+\epsilon_3))\ldots)(1+\epsilon_n) \\ &= \sum_k x_k y_k (1+\mu_k) \end{aligned}$$

z $|\mu_k| \leq n\nu + O(\nu^2)$. Przyjmując, że $\hat{y}_k = y_k(1+\mu_k)$ otrzymujemy, że

$$\tilde{s} = fl(s) = f(x, \hat{y})$$

z ($x = \hat{x}$)

$$\begin{aligned} \sum_k (|x_k - \hat{x}_k| + |y_k - \hat{y}_k|) &= \sum_k |y_k - \hat{y}_k| \leq \sum_k |\mu_k||y_k| \\ &\leq n * \nu(\sum_k (|y_k| + |x_k|)) + O(\nu^2) \end{aligned}$$

Biorąc $K_1 = n$ i $K_2 = 0$ otrzymujemy numeryczną poprawność algorytmu.

Zad. 4 Rozwiązanie tego zadania wynika z poprzedniego zadania czyli z numerycznej poprawności standardowej metody obliczania iloczynu skalarnego dwóch wektorów, przy czym co ważne wynik jest zaburzonym wynikiem dla zaburzonego tylko jednego z wektorów. ponieważ $y_k = (Ax)_k = x^T w_k$ dla w_k k-tego wiersza macierzy A. A nasz algorytm obliczania y_k polega na obliczaniu standardowym algorytmem iloczynów skalarnych x z kolejnymi wierszami A. Tak więc wystarczy skorzystać z tego że powtarzamy algorytm NP dla każdej składowej czyli w jego wyniku dostaniemy $fl(y_k) = x^T \hat{w}_k$ dla $\hat{w}_k$ zaburzonego k-tego wiersza. Za normę danych można przyjąć sumę modułów wszystkich danych.