Kombinatoryczne Algorytmy dla Algebraicznych Problemów Związanych z Macierzami

Autoreferat rozprawy doktorskiej

Anna Urbańska

Instytut Informatyki
Uniwersytet Warszawski

7 stycznia 2009

Wprowadzenie

Metoda kombinatoryczna liczenia problemów macierzowych to sprowadzanie problemów natury numerycznej do prostych problemów teorii grafów i kombinatoryki. Szczególną specyfiką tej metody jest unikanie dzielenia. Wyeliminowanie dzielenia jest istotnym czynnikiem podczas pracy z pierścieniem przemiennym, który nie jest ciałem, na przykład, kiedy elementy pierścienia są liczbami całkowitymi, wielomianami czy też jeszcze bardziej skomplikowanymi wyrażeniami. Takie obliczenia pojawiają się w wielu kombinatorycznych problemach [9].

W rozprawie prezentujemy zastosowanie metod kombinatorycznych w konstrukcji efektywnych algorytmów nie wykonujących operacji dzielenia rozwiązujących takie problemy algebry liniowej, jak: obliczanie wyznacznika, Pfaffianu, współczynników wielomianu charakterystycznego oraz macierzy dopełnień algebraicznych. Konstruowane przez nas algorytmy wykorzystują jako procedury operacje grafowe. Dzięki temu zyskują na czytelności.

Tło

Pierwsze próby zastosowania technik kombinatorycznych do dowodzenia poprawności lub tworzenia nowych algorytmów obliczających funkcje macierzowe zostały podjęte przez Valianta [26] w 1992 roku. Valiant analizował algorytm Samuelsona obliczania współczynników wielomianu charakterystycznego i zinterpretował te obliczenia w sposób kombinato-

ryczny. Przedstawił on kombinatoryczne twierdzenie posługując się terminologią zamkniętych ścieżek w grafie, poprawność którego wynikała wprost z poprawności algorytmu Samuelsona. Zainspirowani tym oraz czysto kombinatorycznym oraz wyjątkowo eleganckim dowodem twierdzenia Cayleya-Hamiltona podanym przez Rutherforda [19], Mahajan i Vinay [16] zaprezentowali pierwszy kombinatoryczny algorytm obliczania wielomianu charakterystycznego. Dowód poprawności ich algorytmu odwoływał się jedynie do argumentów kombinatorycznych bez korzystania z algebry liniowej lub arytmetyki na wielomianach.

Główne wyniki rozprawy

Wyznacznik. Jednym z klasycznych problemów algebry liniowej jest problem obliczania wyznacznika. *Wyznacznik* macierzy kwadratowej A rozmiaru $n \times n$, $det(A)$, zdefiniowany jest jako

$$det(A) = (-1)^n \cdot \sum_{\sigma} sgn(\sigma) \cdot weight(\sigma, A),$$

gdzie sumowanie jest po wszystkich permutacjach σ grupy permutacji zbioru $\{1, 2, \ldots, n\}$, *znak* permutacji σ, $sgn(\sigma)$, równy jest $(-1)^k$, gdzie k jest liczbą cykli w rozkładzie na cykle permutacji σ, natomiast *waga* permutacji σ wynosi

$$weight(\sigma, A) = A[1, \sigma(1)] \cdot A[2, \sigma(2)] \cdot \ldots \cdot A[n, \sigma(n)].$$

Najbardziej znanym algorytmem obliczania wyznacznika jest algorytm eliminacji Gaussa. Wymaga on $O(n^3)$ (lub $O(n^{2.38})$, jeśli użyty zostanie algorytm szybkiego mnożenia macierzy) operacji dodawania, odejmowania, mnożenia oraz *dzielenia*. Z drugiej strony, definicja wyznacznika jako sumy $n!$ iloczynów pokazuje, że może on zostać obliczony *bez* wykonywania operacji dzielenia.

Mahajan i Vinay w swojej pracy z 1997 roku [15] przedstawili zupełnie nową kombinatoryczną interpretację wyznacznika redukując problem jego obliczania do problemu sumowania ścieżek w pewnym acyklicznym grafie a następnie bazując na tej charakteryzacji podali algorytm obliczania wyznacznika w czasie $O(n^4)$, który nie wykorzystywał operacji dzielenia.

W rozprawie przedstawimy nowe spojrzenie na algorytm Mahajana i Vinaya: relację z pseudo-wielomianowym algorytmem programowania dynamicznego dla problemu plecakowego. Główna faza algorytmu Mahajana i Vinaya może być zinterpretowana jako obliczenia algebraicznej wersji problemu plecakowego, co jest alternatywą dla podejścia opartego na teorii grafów użytego w oryginalnym algorytmie.

Głównym rezultatem będzie pokazanie jak zaimplementować algorytm Mahajana i Vinaya bez korzystania z operacji dzielenia w czasie $\tilde{O}(n^{3.03})$. Zaprezentowany przez nas algorytm posiada bardzo prostą implementację w czasie $O(n^{3.5})$ a jedynym nietrywialnym elementem algorytmu działającego w czasie $\tilde{O}(n^{3.03})$ jest algorytm szybkiego mnożenia macierzy.

Pfaffian. W pracy rozważamy również Pfaffian macierzy skośno-symetrycznych ściśle związany z pojęciem wyznacznika. *Pfaffian* skośno-symetrycznej macierzy A (tzn. $A = -A^T$) o parzystej liczbie n wierszy i kolumn zdefiniowany jest jako

$$Pf(A) = \sum_{\mathcal{M}} sgn(\mathcal{M}) \cdot weight(\mathcal{M}, A),$$

gdzie sumowanie odbywa się po wszystkich skojarzeniach doskonałych $\mathcal{M}$ grafu pełnego K_n. Skojarzenie doskonałe $\mathcal{M}$ zapisujemy jako podział zbioru $\{1, 2, \ldots, n\}$ na $m = n/2$ nieuporządkowanych par

$$\mathcal{M} = \{\{i_1, j_1\}, \{i_2, j_2\}, \ldots, \{i_m, j_m\}\},$$

gdzie $i_k < j_k$ dla każdego $k = 1, 2, \ldots, m$ oraz $i_1 < i_2 < \ldots < i_m$. *Znak* skojarzenia $\mathcal{M}$, $sgn(\mathcal{M})$, definiujemy jako znak permutacji

$$\sigma_{\mathcal{M}} = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & n-1 & n \\ i_1 & j_1 & i_2 & j_2 & \cdots & i_m & j_m \end{pmatrix}.$$

Waga skojarzenia $\mathcal{M}$ jest równa

$$weight(\mathcal{M}, A) = A[i_1, j_1] \cdot A[i_2, j_2] \cdot \ldots \cdot A[i_m, j_m].$$

Pfaffian pojawia się w badaniach nad skojarzeniami w grafach [11]: Pfaffian zorientowanego grafu zdefiniowany jest jako suma przebiegająca po wszystkich możliwych skojarzeniach doskonałych, przy czym każde skojarzenie ma przypisany znak zależny od orientacji. To wskazuje już na podobieństwa do wyznacznika. Gdyby nie obecność znaku w tej definicji Pfaffian można by wykorzystać do wyznaczania liczby skojarzeń doskonałych. Wiadomo, że istnieją specjalne grafy, które można zorientować w taki sposób, aby wszystkie skojarzenia doskonałe miały dodatni znak. To oczywiście oznacza, że żadne skracanie się składników w sumie nie będzie miało miejsca a stąd Pfaffian da dokładnie liczbę skojarzeń doskonałych.

Pfaffian możemy wyznaczyć wykonując procedurę eliminacji dla macierzy skośno-symetrycznych, która jest podobna do eliminacji Gaussa (i wykorzystuje operację dzielenia). Alternatywnie, pierwiastek kwadratowy z wyznacznika daje Pfaffian z dokładnością do znaku. Żadne z tych podejść nie może jednak zostać użyte, jeśli operacje dzielenia i pierwiastkowania będą musiały zostać wyeliminowane.

W naszej pracy korzystając z techniki kombinatorycznej opracowanej przez Mahajana, Subramanya i Vinaya [14] dostajemy algorytm obliczania Pfaffianu działający w czasie $\tilde{O}(n^{3.03})$ i nie wykonujący operacji dzielenia.

Wielomian charakterystyczny. Jak się okazuje metoda obliczania wyznacznika wprowadzona przez Mahajana i Vinaya może być wykorzystana do wyznaczania *wszystkich* współczynników *wielomianu charakterystycznego* macierzy A [16]:

$$\Phi_A(\lambda) = det(\lambda \cdot I_n - A) = c_0 \cdot \lambda^n + c_1 \cdot \lambda^{n-1} + \ldots + c_{n-1} \cdot \lambda + c_n,$$

gdzie I_n oznacza macierz identycznościową rozmiaru $n \times n$.

W rozprawie opisujemy algorytm obliczania wielomianu charakterystycznego który tak jak w przypadku wyznacznika nie wykorzystuje operacji dzielenia i działa w czasie $\tilde{O}(n^{3.03})$.

Macierz dopełnień algebraicznych. Dzięki technice wprowadzonej przez Baura i Strassena [1] pokazującej jak obliczać wszystkie pochodne cząstkowe otrzymujemy, że dowolna metoda obliczania wyznacznika może zostać automatycznie skonwertowana do metody obliczającej macierz dopełnień algebraicznych, przy czym złożoność wzrasta o nie więcej niż stały czynnik. *Macierz dopełnień algebraicznych*, $adj(A)$, jest macierzą rozmiaru $n \times n$ taką, że

$$adj(A)[i,j] = (-1)^{i+j} det(A^{i,j}),$$

gdzie $A^{i,j}$ oznacza macierz rozmiaru $(n-1) \times (n-1)$ powstałą z macierzy A przez usunięcie wiersza o numerze i i kolumny o numerze j.

Dostajemy tym samym nie wykonujący dzielenia algorytm obliczania macierzy dopełnień algebraicznych działający w czasie $\tilde{O}(n^{3.03})$.

Grafy planarne. Problem obliczania wyznacznika wydaje się być znacznie prostszy dla macierzy określonych przez grafy planarne, tzn. grafy, które można narysować na płaszczyźnie tak, by krzywe obrazujące krawędzie grafu nie przecinały się ze sobą.

Na samym początku takie macierze są rzadkie, mają one jedynie $O(n)$ niezerowych współrzędnych. Jednak możemy powiedzieć znacznie więcej. Lipton, Rose oraz Tarjan [12] pokazali, że istnienie rodziny $O(\sqrt{n})$-separatorów dla tej klasy grafów daje możliwość wykonania eliminacji Gaussa w czasie $O(n^{3/2})$ (lub $O(n^{1.19})$, jeśli użyty zostanie algorytm szybkiego mnożenia macierzy). Jednak działanie tego algorytmu ciągle wymaga dzielenia.

W naszej pracy prezentujemy algorytmy obliczania wyznacznika, wielomianu charakterystycznego oraz macierzy dopełnień algebraicznych dla przypadku grafów planarnych działające w czasie $O(n^{2.5})$ i nie wykonujące operacji dzielenia.

Inne zastosowania. W rozprawie pokazujemy również jak możemy wykorzystać nasze wyniki do implementacji znanych algorytmów obliczania wyznacznika: algorytmu Samuelsona oraz algorytmu Chistova w czasie $\tilde{O}(n^{3.03})$.

Uwagi końcowe

Nowe wyniki naukowe, przedstawione w rozprawie, zostały również zawarte w następujących artykułach:

- A. Urbańska, *Faster Combinatorial Algorithms for Determinant and Pfaffian,* ISAAC 2007, LNCS 4835 (2007) 599–608 [24]
- A. Urbańska, *Faster Combinatorial Algorithms for Determinant and Pfaffian,* Praca przyjęta do publikacji w czasopiśmie Algorithmica, publikacja elektroniczna dostępna na stronie: http://dx.doi.org/10.1007/s00453-008-9240-9 (2008) [25]
- A. Urbańska, *Application of Graph Separators to the Efficient Division-Free Computation of Determinant,* Praca przyjęta na konferencję SOFSEM 2009 [23]

Literatura

[1] W. Baur and V. Strassen, *The Complexity of Partial Derivatives,* Theoretical Comput. Sci. **22** (1983) 317–330

[2] A. L. Chistov, *Fast Parallel Calculation of the Rank of Matrices over a Field of Arbitrary Characteristic,* In Proc Int. Conf. Foundations of Computation Theory, LNCS 199 Springer (1985) 63–69

[3] D. Coppersmith, *Rectangular Matrix Multiplication Revisited,* J. Complex. **13(1)** (1997) 42–49

[4] D. Coppersmith and S. Winograd, *Matrix Multiplication via Arithmetic Progressions,* In Proceedings of the nineteenth Annual ACM Conference on Theory of Computing (1987) 1–6

[5] P. Doubilet, *On the Foundations of Combinatorial Theory,* VII: symmetric functions through the theory of distribution and occupancy, Stud. Appl. Math. Vol. LI No. 4 (1972) 377.396, reprinted in: Gian-Carlo Rota on Combinatorics, J.P. Kung Ed., Birkh¨auser (1995) 403–422.

[6] D. Fadeev and V. Fadeeva, *Computational Methods in Linear Algebra,* Freeman, San Francisco (1963)

[7] J. Geelen, *An Algebraic Matching Theory,* Combinatorica **20** (2000) 61–70

[8] M. Karpiński and W. Rytter, *Fast Parallel Algorithms for Graph Matching Problems,* Oxford University Press, Oxford (1998)

[9] C. Krattenthaler, *Determinant Calculus,* Séminaire Lotharingien de Combinatoire B42 (1999) 67

[10] L. Lovász, *On Determinants, Matchings and Random Algorithms,* In L. Budach, editor, Fundamentals of Computation Theory, Akademie-Verlag (1979) 565–574

[11] L. Lovász and M. Plummer, *Matching Theory,* Ann. Discr. Math, Vol. 29. North-Holland Mathematics Studies, Vol. 121, Amsterdam (1986)

[12] R. J. Lipton, D. J. Rose and R. E. Tarjan, *Generalized Nested Dissection,* SIAM J. Num. Anal. **16** (1979) 346–358

[13] R. J. Lipton and R. E. Tarjan, *A Separator Theorem for Planar Graphs,* SIAM J. Applied Math. (1979) 177–189

[14] M. Mahajan, P. Subramanya and V. Vinay, *A Combinatorial Algorithm for Pfaffians,* In Computing and Combinatorics. Proc. Fifth Annual International Conference. (COCOON '99), Tokyo, July 1999, ed. Takao Asano et al. Lecture Notes Computer Science 1627, Springer-Verlag (1999) 134–143

[15] M. Mahajan and V. Vinay, *A Combinatorial Algorithm for the Determinant,* In Proceedings of the Eight Annual ACM-SIAM Symposium on Discrete Algorithms, SODA97

[16] M. Mahajan and V. Vinay, *Determinant: Combinatorics, Algorithms and Complexity,* Chicago Journal of Theoretical Computer Science **5** (1997)

[17] M. Mahajan and V. Vinay, *Determinant: Old algorithms, New Insights,* SIAM J. Discrete Math. **12** (1999) 474–490

[18] G. Rote, *Division-Free Algorithms for the Determinant and the Pfaffian: Algebraic and Combinatorial Approaches,* In H. Alt, editor, Computational Discrete Mathematics: Advanced Lectures, volume LNCS 2122, Springer (2001) 119–135

[19] D. E. Rutherford, *The Cayley-Hamilton Theorem for Semi-Rings,* Proc. Roy. Soc. Edinburgh, Sect. A **66** (1964) 211–215

[20] P. A. Samuelson, *A Method of Determining Explicitly the Coefficients of the Characteristic Polynomial,* Ann. Math. Stat. **13** (1942) 424–429

[21] V. Strassen, *Vermeidung von Divisionen,* Journal of Reine U. Angew Math. **264** (1973) 182–202

[22] H. Straubing, *A Combinatorial Proof of the Cayley-Hamilton Theorem,* Discrete Math. **43** (1983) 273–279

[23] A. Urbańska, *Application of Graph Separators to the Efficient Division-Free Computation of Determinant,* Praca przyjęta na konferencję SOFSEM 2009

[24] A. Urbańska, *Faster Combinatorial Algorithms for Determinant and Pfaffian,* ISAAC 2007, LNCS 4835 (2007) 599–608

[25] A. Urbańska, *Faster Combinatorial Algorithms for Determinant and Pfaffian,* Praca przyjęta do publikacji w czasopiśmie Algorithmica, publikacja elektroniczna dostępna na stronie: http://dx.doi.org/10.1007/s00453-008-9240-9 (2008)

[26] L. G. Valiant, *Why is Boolean Complexity Theory Difficult?* In: Boolean Function Complexity, ed. M. S. Paterson, LMS Lecture Notes Series, Vol. 169, Cambridge Univ. Press (1992) 84–94

[27] L. G. Valiant, *The Complexity of Computing the Permanent,* Theoretical Computer Science **8** (1979) 189–201

[28] V. Vinay, *Counting Auxiliary Pushdown Automata and Semi-Unbounded Arithmetic Circuits,* In Proc. 6th Structure in Complexity Theory Conference (1991) 270–284

[29] D. B. Wilson, *Determinant Algorithms for Random Planar Structures,* In Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics (1997) 258–267

(1985) 61–72