Structures de déplacement pour les matrices p-Vandermonde confluentes et éliminations de Gauss avec pivotement partiel

Lamine Melkemi

Laboratoire de l'Informatique du Parallélisme

École Normale Supérieure de Lyon
Unité de recherche associée au CNRS n° 1398

Structures de déplacement pour les matrices p-Vandermonde confluentes et éliminations de Gauss avec pivotement partiel

Lamine Melkemi

May 1998

Research Report N° 98-23

École Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon Cedex 07, France
Téléphone : +33(0)4.72.72.80.00
Télécopieur : +33(0)4.72.72.80.80
Adresse électronique : `lip@ens-lyon.fr`

Structures de déplacement pour les matrices p-Vandermonde confluentes et éliminations de Gauss avec pivotement partiel

Lamine Melkemi

May 1998

Abstract

In this paper, we show that the useful displacement structures constructed for polynomial Vandermonde matrices (see [8] in particular) can be naturally extended to confluent polynomial Vandermonde matrices. This result was made possible by virtue of the fact (recently established by the author in [10]) that confluent Vandermonde matrices belong favorably to the class of structured matrices. In the context of the displacement structure theory, it is well known that once an acceptable displacement structure is established for a matrix, one may naturally expect interesting applications regarding its numerical implementation.

Keywords: Confluent Vandermonde matrix, confluent polynomial Vandermonde matrix, displacement structure.

Résumé

Dans cet article, nous construisons des structures de déplacement utiles pour les matrices p-Vandermonde confluentes, généralisant ainsi les résultats connus dans le cas des matrices p-Vandermonde [8]. L'idée principale repose sur un résultat récent établi par l'auteur [10] affirmant que les matrices de Vandermonde confluentes font partie, tout comme les matrices de Vandermonde, de la classe des matrices structurées.

Mots-clés: matrice de Vandermonde confluente,matrice $p-$Vandermonde confluente , structure de déplacement .

1.Introduction. Soient $x_0,\ x_1,\ ..., x_{n-1} \in \mathbb{R}$ n nombres réels deux à deux distincts, et soit

$$p = \{p_0(x), p_1(x)..., p_{m-1}(x)\}$$

une base de l'espace $\mathbb{R}_{m-1}[x]$ des polynômes de degrè inférieur ou égal à $m-1$. Parmi ces bases, on distingue la base canonique qu'on note

$$c = \{1,\ x,\ x^2, \ldots, x^{m-1}\}$$

Considérons alors la matrice suivante du type $m \times dn$:

$$V_p = [f(x_0)\ \ f'(x_0) \ldots f^{(d-1)}(x_0) \ldots f(x_{n-1})\ \ f'(x_{n-1}) \ldots f^{(d-1)}(x_{n-1})] \tag{1.1}$$

où $f(x)$ est la fonction vectorielle:

$$f(x) = (p_0(x), p_1(x), \ldots, p_{m-1}(x))^T$$

et $f'(x),\ f^{(2)}(x), ..., f^{(d-1)}(x)$ sont les dérivées successives de $f(x)$.

Définition 1.1. La matrice V_c donnée par (1.1) en remplaçant p par la base canonique c est appelée matrice de Vandermonde confluente. En général, V_p est appelée matrice $p-$Vandermonde confluente. □

Dans cette note, on considére en particulier les bases p de $\mathbb{R}_{m-1}[x]$ vérifiant des relations récurrentes du type suivant:

$$p_0(x) = 1,\ p_i(x) = 0,\ i < 0,\ p_i(x) = \alpha x p_{i-1}(x) + \sum_{j=1}^{k-1} \gamma_j p_{i-j}(x) \tag{1.2}$$

Notre propos consiste alors à montrer que la matrice V_p pour une base p vérifiant (1.2) satisfait une structure de déplacement acceptable. C'est à dire qu'il existe deux 'simples' matrices (que nous préciserons plus loin) L et U telles que:

$$LV_p - V_pU \ = \ e_m y^T$$

où $y \in \mathbb{R}^{dn}$ et $\{e_1, \ldots, e_m\}$ désigne la base canonique de $\mathbb{R}^m$. On renvoit le lecteur à [6], [11], [1] et au récent article de synthèse [9] pour plus de détails concernant les structures de déplacement. Dans ce contexte, on invoque souvent la matrice de déplacement du type $m \times m$:

$$Z = \begin{bmatrix} 0 & & \ldots & \ldots & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & 0 \\ 0 & \ldots & 0 & 1 & 0 \end{bmatrix}$$

Le résultat que nous établissons repose sur des techniques similaires à celles utilisées dans [7], et exploite un résultat dû à [10] qui montre qu'une structure de déplacement simple caractérisant bien les matrices de Vandermonde confluentes peut être construite:

Théorème 1.1.([10]) Soit V_c la matrice de Vandermonde confluente donnée par (1.1) où p est remplacée par c. Alors

$$Z^T V_c - V_c D = e_m y^T \tag{1.3}$$

où Z est la matrice de déplacement d'ordre m et $D = diag(B_0, \ldots, B_{n-1})$ avec

$$B_i = B(x_i) = \begin{bmatrix} x_i & 1 & 0 & \ldots & 0 \\ 0 & x_i & 2 & \ldots & 0 \\ 0 & 0 & x_i & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & d-1 \\ 0 & 0 & \ldots & \ldots & x_i \end{bmatrix}.$$

Idées de la démonstration. Notons par W la matrice de Vandermonde par blocs suivante:

$$W = \begin{bmatrix} I_d & I_d & \ldots & I_d \\ B_0 & B_1 & \ldots & B_{n-1} \\ \vdots & \vdots & & \vdots \\ B_0^{m-1} & B_1^{m-1} & \ldots & B_{n-1}^{m-1} \end{bmatrix}$$

L'idée clé consiste à montrer que la matrice de Vandermonde confluente V_c est plongée dans W, c'est à dire qu'il existe une matrice de permutation P telle que $PW = \begin{bmatrix} V_c \\ X \end{bmatrix}$. Ceci découle de l'observation que la première ligne de la matrice $B^k(x)$ (rappelons que $B(x_i) = B_i$) est formée de x^k et de ses dérivées successives kx^{k-1},... Enfin le théorème s'ensuit en appliquant ce plongement à la structure de déplacement de W. $\square$

2. Structure de déplacement pour V_p. Soit p une base de $\mathbb{R}_{m-1}[x]$ vérifiant la relation récurrente (1.2), et soit $A = (a_{ij})_{0 \le i,j < m}$ la matrice du type $m \times m$ définie comme suit:

$$p_i(x) \; = \; \sum_{j=0}^{m-1} a_{ij} x^j$$

Alors, il n'est pas difficile de montrer que:

$$V_p \; = \; A V_c \tag{2.1}$$

Par ailleurs, comme la base p satisfait la relation récurrente (1.2), il est légitime d'espérer que les éléments a_{ij} de A soient liés entre eux par une relation récurrente analogue. En effet, en identifiant les coefficients de x^j dans les deux membres de la relation présentée dans (1.2), on montre facilement que:

$$a_{ij} = \alpha a_{i-1,j-1} + \textstyle\sum_{s=1}^{k-1} \gamma_s a_{i-s,j} \tag{2.2}$$

Ces relations sont à la base de notre résultat. En indroduisant la matrice de déplacement Z et en observant qu'intuitivement, a_{ij} correspond à A, $a_{i-1,j-1}$ correspond à ZAZ^T et $a_{i-s,j}$ correspond à $Z^s A$ (par convention $Z^0 = I_m$), on peut énoncer le lemme suivant où une structure de déplacement de A est construite:

Lemme 2.1. Supposons $\alpha \neq 0$. Il existe un vecteur $g \in \mathbb{R}^{dn}$ pour lequel la matrice A vérifie la structure de déplacement suivante:

$$\frac{1}{\alpha}(Z^T - \textstyle\sum_{s=1}^{k-1} \gamma_s Z^{s-1})A - AZ^T \;=\; e_m g^T. \tag{2.3}$$

Démonstration. En examinant de près les relations (1.2) et (2.2), il n'est pas difficile de montrer que:

$$A = \alpha Z A Z^T + \sum_{s=1}^{k-1} \gamma_s Z^s A + e_1 e_1^T.$$

En multipliant alors les deux membres par Z^T et en observant que $Z^T e_1 = 0$ et $Z^T Z = I_m - e_m e_m^T$, on obtient:

$$Z^T A = \alpha(I_m - e_m e_m^T) A Z^T + \sum_{s=1}^{k-1} \gamma_s (I_m - e_m e_m^T) Z^{s-1} A$$

ce qui peut s'écrire encore de la manière suivante:

$$(Z^T - \sum_{s=1}^{k-1} \gamma_s Z^{s-1})A - \alpha A Z^T = e_m g'^T$$

où $g'^T = -\alpha e_m^T A Z^T - \sum_{s=1}^{k-1} \gamma_s e_m^T Z^{s-1} A$. Le lemme s'ensuit en divisant les deux membres par α et en posant $g = \frac{1}{\alpha} g'$. $\square$

Grâce à ce lemme, on est en mesure d'énoncer le résultat principal:

Théorème 2.2. Soit p une base de $\mathbb{R}_{m-1}[x]$ vérifiant la relation récurrente (1.2) avec $\alpha \neq 0$. Alors, la matrice V_p donnée par (1.1) vérifie la structure de déplacement suivante:

$$\frac{1}{\alpha}(Z^T - \textstyle\sum_{s=1}^{k-1} \gamma_s Z^{s-1})V_p - V_p D \;=\; e_m z^T. \tag{2.4}$$

où $z = \alpha^{m-1} y + V_c^T g$. y et D étant introduits dans le théorème 1.1.

Démonstration. Au vu de la relation (1.2) vérifiée par la base p, on peut affirmer que $deg p_i(x) = i,\ 0 \leq i < m$. En conséquence, la matrice A est triangulaire inférieure, et il est aisé de montrer que $a_{ii} = \alpha^i$. D'après le lemme précédent, on peut écrire

$$\frac{1}{\alpha}(Z^T - \sum_{s=1}^{k-1} \gamma_s Z^{s-1})V_p = (A Z^T + e_m g^T) V_c$$

(on a utilisé ici le fait que $V_p = A V_c$). En appliquant maintenant le théorème 1.1, on obtient:

$$\frac{1}{\alpha}(Z^T - \sum_{s=1}^{k-1} \gamma_s Z^{s-1})V_p = A(V_c D + e_m y^T) + e_m g^T V_c$$

Le théorème découle directement du fait que A est triangulaire inférieure avec $a_{ii} = \alpha^i$, et par conséquent $Ae_m = \alpha^{m-1} e_m$. $\square$

Posons $u = \frac{1}{\alpha}(0, ..., 0, \gamma_{k-1}, ..., \gamma_2)^T \in \mathbb{R}^m$, et remarquons que:

$$C = \tfrac{1}{\alpha}(Z^T - \textstyle\sum_{s=1}^{k-1} \gamma_s Z^{s-1}) + \tfrac{1}{\alpha} e_m e_1^T - \textstyle\sum_{s=0}^{k-2} e_{s+1} u^T Z^s \tag{2.5}$$

est une matrice circulante si bien qu'on peut montrer facilement qu'il existe k vecteurs $w_1, \ldots, w_k \in \mathbb{R}^{dn}$ pour lesquels la structure de déplacement (2.4) devient:

$$CV_p - V_p D = [e_1 \; ... \; e_{k-1} \; e_m][w_1 \; ... \; w_k]^T \tag{2.6}$$

D'autre part, puisque C est circulante, on peut déterminer directement ses valeurs et vecteurs propres. Plus précisément, on a:

$$FCF^H = \Sigma = diag(\sigma_1, \ldots, \sigma_m) \tag{2.7}$$

où $F = \frac{1}{\sqrt{n}}(e^{\frac{2\pi i}{m} kj})_{0 \leq k,j < m}$ est la transformée discrète normalisée de Fourier (TDFn), et $(\sigma_1, \ldots, \sigma_m)^T$ est la TDFn de la première colonne de C. Ces observations donnent lieu au résultat suivant où une structure de déplacement pour $\hat{V}_p = FV_p$ est déduite:

Theorem 2.3. Soit Σ la matrice diagonale définie dans (2.7), et posons $\hat{V}_p = FV_p$, F étant la TDFn présentée dans (2.7). Alors

$$\Sigma \hat{V}_p - \hat{V}_p D = [f_1 \; ... \; f_{k-1} \; f_m][w_1 \; ... \; w_k]^T \tag{2.8}$$

où $f_j = Fe_j$.

Démonstration. Le théorème s'en déduit directement en pré-multipliant les deux membres de la structure (2.6) par F, et en observant que $FCV_p = \Sigma \hat{V}_p$. $\square$

3. Eliminations de Gauss rapides avec pivotement partiel.

Dans cette section, l'objet est de présenter une méthode de triangularisation rapide appliquée à V_c et plus généralement à V_p, qu'on appelle éliminations de Gauss rapides avec pivotement partiel par analogie au procédé classique des éliminations de Gauss avec pivotement partiel (EGPP) utilisé pour résoudre les systèmes linéaires. On renvoit le lecteur à [3] , entre autres, où la méthode EGPP est soigneusement décrite et analysée. En ce qui nous concerne, on rappelle juste la première étape de EGPP qui représente bien les étapes d'après. Soit à appliquer cette méthode à une matrice régulière M du type $m \times m$. Alors dans la première étape de EGPP, on procéde en deux temps: la matrice M est d'abord multipliée par une matrice de permutation P de sorte que

$$piv = (PM)_{1,1} = Max\{(PM)_{i,1}; 1 \leq i \leq m\}.$$

Ensuite on multiplie PM par la matrice $J = I_m - \frac{1}{piv} g e_1^T$, où $g = PMe_1$ est la première colonne de PM. ALors, on vérifie facilement que:

$$JPM = \begin{bmatrix} piv & * \\ 0 & N \end{bmatrix}. \tag{3.1}$$

Enfin du point de vue algorithmique, le but est de calculer et de recouvrir la première ligne de PM. La seconde étape consiste à appliquer cette première étape à la matrice N; et de cette façon même, on déduit les étapes $3, ..., m$ de la méthode EGPP. Il est bien connu que cette méthode utilise $O(n^3)$ multiplications; ce qui est toutefois un inconvénient une fois la méthode est appliquée à des matrices ayant des formes particulières. Dans ce qui suit, on procéde à adapter les EGPP dans le cas où $M = V_c$.

Définition 3.1. La matrice N définie dans (3.1) est appelée le complément de Schur de la matrice PM. On note par $\mathcal{S}(R)$ le complément de Schur (s'il existe) de la matrice R. Ainsi $N = \mathcal{S}(PM)$. □

Dans notre contexte, l'introduction de cette définition a un double intérêt. D'un coté, le complément de Schur est, comme on vient de le voir, la pierre de fondation dans le procédé des EGPP. De l'autre, on peut dire vaguement que l'opérateur $\mathcal{S}$ du complément de Schur est en général invariant par les structures de déplacement "acceptables". Avant d'exploiter ces observations, nous tenons à faire remarquer d'abord que les EGPP rapides seront appliquées plutôt à la TDFn $\hat{V}_c = FV_c$ de V_c, qui, elle aussi, est dotée d'une structure de déplacement que nous présentons dans le résultat suivant:

Théorème 3.1. La TDFn $\hat{V}_c = FV_c$ de V_c satisfait la structure de déplacement suivante:

$$\Omega \hat{V}_c - \hat{V}_c D = f w^T \tag{3.2}$$

où $f = Fe_m$, $w = y + V_c^T e_1$ et $\Omega = diag(1, e^{-\frac{2\pi i}{m}}, ..., e^{-\frac{2\pi i}{m}(n-1)})$.

Démonstration. En ajoutant $e_m e_1^T V_c$ dans les deux membres de la structure (1.3), on obtient directement:

$$(Z^T + e_m e_1^T)V_c - V_c D = e_m(y^T + e_m e_1^T V_c)$$

on observe par ailleurs que $(Z^T + e_m e_1^T)$ est une matrice circulante et que $F(Z^T + e_m e_1^T)F^H = \Omega$ de sorte que si l'on multiplie cette structure par F, la relation (3.2) s'ensuit directement. □

Il importe de remarquer dans ce résultat et dans le théorme 2.3, que l'opérateur de Sylvester est du même genre. Appliqué à une matrice, disons R (dans notre cas $R = V_c$ ou V_p), un tel opérateur est de la forme

$$Sylv_{\Delta,B}(R) = \Delta R - RB \tag{3.3}$$

où Δ est une matrice diagonale et B une matrice bidiagonale. Pour appliquer les EGPP rapides à $\hat{V}_c$, notons par $\mathcal{CV}$ l'ensemble des matrices R du type $s \times s$

($s \in \mathbb{N}^*$) pour lesquelles il exite deux matrices diagonale et bidiagonale Δ, $B \in \mathbb{R}^{s\times s}$ respectivement et deux vecteurs $u, v \in \mathbb{R}^s$ tels que:

$$Sylv_{\Delta,B}(R) = uv^T$$

où $Sylv_{\Delta,B}$ est définie dans (3.3). On voit immédiatement que $\hat{V}_c \in \mathcal{CV}$.

L'approche rapide des EGPP appliquée à $\hat{V}_c$ est rendue possible grâce aux deux résultats suivants dûs à [1].

Observation 3.2. Soit R une matrice du type $s \times s$ et appartenant à $\mathcal{CV}$. Alors pour toute matrice de permutation $P \in \mathbb{R}^{s\times s}$, on a $PR \in \mathcal{CV}$.

Démonstration. Elle découle directement du fait que si Δ est une matrice diagonale, alors $P\Delta P^T$ l'est aussi. □

Théorème 3.3.([1], [2]) Soit R une matrice du type $s \times s$ ($s \geq 2$) appartenant à $\mathcal{CV}$; c'est à dire que:

$$\Delta R - RB = uv^T$$

Ecrivons par blocs:

$$R = \begin{bmatrix} r_{11} & h^T \\ g & R' \end{bmatrix}, \quad \Delta = \begin{bmatrix} \delta & 0 \\ 0 & \Delta' \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & * \\ 0 & B' \end{bmatrix}$$

et supposons que $r_{11} \neq 0$ de sorte que le complément de Schur de R existe. Alors le complément de Schur $\mathcal{S}(R)$ de R appartient à $\mathcal{CV}$. De plus, on a:

$$\Delta'\mathcal{S}(R) - \mathcal{S}(R)B' = u'v'^T$$

avec

$$\begin{bmatrix} 0 \\ u' \end{bmatrix} = u - \frac{u_1}{r_{11}} \begin{bmatrix} r_{11} \\ g \end{bmatrix}, \quad [0 \quad v'^T] = v^T - \frac{v_1}{r_{11}}[r_{11} \quad h^T]$$

où u_1 et v_1 sont respectivement les premières composantes de u et de v. □

Considérons de nouveau la première étape des EGPP où l'on avait principalement obtenu la relation (3.1). Dans le cas où $M = V_c$ (ou plus généralement $M \in \mathcal{CV}$) alors on observe d'abord que $PM \in \mathcal{CV}$ d'après l'observation 3.2; ensuite compte tenu du théorème 3.3, $N \in \mathcal{CV}$ puisque $N = S(PM)$. D'autant plus, le théorème 3.3 fournit les formules permettant de construire la structure de N. Pour que la description de la première étape des EGPP rapides soit aussi complète que possible, nous n'avons qu'à calculer la première ligne de $S(PM)$ en vu de recouvrir la matrice triangulaire supérieure; ce qui peut se faire facilement en utisant O(n) opérations seulement.

4.Conclusion. Dans cet article, nous avons montré que les matrices p-Vandermonde confluentes satisfont des structures de déplacement acceptables comparablement à celles construites pour les matrices de Toeplitz ou de Cauchy [6], [4]. Dans ce contexte, on sait qu'une fois une telle structure est établie, différentes applications numériques intéressantes en découlent directement. Par exemple, comme on vient de le préciser à la troisième section, grâce aux structures de déplacement (2.8) et (3.2) (où il importe de voir que Σ et Ω sont diagonales), on peut appliquer, avec succès, aux matrices $\hat{V}_c$ et $\hat{V}_p$ par conséquent V_c et V_p la méthode rapide de Gauss avec pivotement partiel développée par Gohberg et al [1]. Contrairement à la méthode classique de Gauss avec pivotement qui exige $O(n^3)$ opérations, la complexité de la méthode rapide dans [1] est de $O((dn)^2)$ pour V_c et $O(k(dn)^2)$ pour V_p.

De fait, de telles structures ont été déjà développées dans la litérature dans le cas non confluent (voir [5], [7], [8]). Grâce au théorème 1.1 où une structure de déplacement pour les matrices de Vandermonde confluentes est conçue [10], nous avons montré, dans cette note, que les structures déjà construites pour les matrices p-Vandermonde [8] peuvent être naturellement étendues au cas confluent.

References

[1] I. Gohberg, T. Kailath and V.Olshevsky Fast Gaussian elimination with partial pivoting for matrices with displacement structure *Math. Comput.* 64(1995) pp.1557-1576

[2] G.H. Golub and V. Olshevsky Pivoting for structured matrices with applications *preprint,* 1997

[3] G.H. Golub and C.F. VanLoan Matrix Computations John Hopkins U.P.Baltimore. Third edition 1996

[4] G. Heinig and K. Rost Algebraic methods for Toeplitz-like matrices and operators *Operator Theory, vol.13, Birkhauser, Basel,* 1984

[5] G. Heinig, W. Hoppe and K. Rost Structured matrices in interpolation and approximation problems Wissenschaftl.Zeitsschrift der TU Karl-Marx-Stadt 31, 2(1989) pp. 196-202

[6] T. Kailath, S.Y. Kung and M. Morf Displacement ranks of matrices and linear equations *J.Math.Anal.Appl.* 68(1979) pp.395-407

[7] T. Kailath and V. Olshevsky Dispalcement structure approach to Chebyshev-Vandermonde and related matrices *Integral Equations and Operator Theory* 22(1995) pp. 65-92

[8] T. Kailath and V. Olshevsky Dispalcement structure approach to polynomial Vandermonde and related matrices *Lin.ALg.Appli.* 261(1997) pp.49-90

[9] T. Kailath and A.H. Sayed Displacement structure: theory and applications *SIAM review* 73(1995) pp. 297-386

[10] L. Melkemi Displacement structure for confluent Vandermonde matrices and applications *Soumis à SIAM J.Matrix Anal.Appli.*

[11] V. Pan On computations with structured matrices *Math.Comp.* 55(1990) pp.179-190