ISSN 1064–5624, Doklady Mathematics, 2008, Vol. 77, No. 2, pp. 1–4. © Pleiades Publishing, Ltd., 2008.
Original Russian Text © V.P. Il'in, 2008, published in Doklady Akademii Nauk, 2008, Vol. 419, No. 3, pp. 000–000.

MATHEMATICS

Conjugate and Semiconjugate Direction Methods with Preconditioning Projectors

V. P. Il'in

Presented by Academician M.M. Lavrent'ev September 28, 2007

Received October 29, 2007

DOI: 10.1134/S.06456240802@@@@

Consider the system of linear algebraic equations (SLAE)

$$Au = f, \quad u = \{u_i\}, \quad f_i = \{f_i\} \in \mathbf{R}^N, \quad A = \{a_{i,j}\} \in \mathbf{R}^{N,N}, \tag{1}$$

for which there is a converging stationary iterative process

$$u^{n+1} = Bu^n + g, \quad u^n \xrightarrow[n\to\infty]{} u, \quad g = (I-B)A^{-1}f \tag{2}$$

with a transition matrix B that has the eigenvalues $\lambda_q(B)$ and the spectral radius $\rho = \max_q\{|\lambda_q(B)|\} < 1$. Then the vector u solves the system

$$\tilde{A}u \equiv (I-B)u = g, \tag{3}$$

where I is the identity matrix and $\tilde{A}$ is the preconditioned matrix with respect to A. If $\tilde{A}$ is a symmetric positive definite (SPD) matrix, then it has the spectral condition number

$$\kappa = \|\tilde{A}\|_2\|\tilde{A}^{-1}\|_2 = \frac{1+\rho}{1-\rho} \tag{4}$$

and SLAE (3) can be solved by an iterative conjugate gradient method (see [1–4]):

$$r^0 = g - \tilde{A}u, \quad p^0 = r^0; \quad n = 0, 1, \ldots;$$
$$u^{n+1} = u^n + \alpha_n^{(s)}p^n, \quad r^{n+1} = r^n - \alpha_n^{(s)}\tilde{A}p^n, \tag{5}$$
$$p^{n+1} = r^{n+1} + \beta_n^{(s)}p^n, \quad s = 0, 1,$$

which is optimal in the Krylov subspaces $\mathcal{H}_{n+1}(r^0, \tilde{A}) = \operatorname{span}\{p^0, \tilde{A}p^0, \ldots, \tilde{A}^n p^0\}$. In the conjugate gradient (CG) and conjugate residual (CR) methods, the iterative parameters $\alpha_n^{(s)}$ and $\beta_n^{(s)}$ in (5) are defined as

$$\alpha_n^{(s)} = \frac{(\tilde{A}^s r^n, r^n)}{(\tilde{A}p^n, \tilde{A}^s p^n)}, \quad \beta_n^{(s)} = \frac{(\tilde{A}^s r^{n+1}, r^{n+1})}{(\tilde{A}^s r^n, r^n)}. \tag{6}$$

Here, $s = 0, 1$ for CG and CR, respectively. These algorithms generate the residual r^n and direction p^n vectors with the orthogonality properties

$$(\tilde{A}^s r^n, r^k) = (\tilde{A}^s r^n, r^n)\delta_{n,k},$$
$$(\tilde{A}p^n, \tilde{A}^s p^k) = (\tilde{A}p^n, \tilde{A}^s p^n)\delta_{n,k} \tag{7}$$

and minimize the functionals $\Phi_n^{(s)}(r^n) = (\tilde{A}^{s-1}r^n, r^n)$ $(s = 0, 1)$ in Krylov subspaces. To reduce these functionals according to the condition $\dfrac{\Phi_n^{(s)}(r^n)}{\Phi_0^{(s)}(r^0)} \le \varepsilon^2 < 1$, the required number of iteration is estimated as

$$n(\varepsilon) \le 1 + \left(\ln\frac{1+\sqrt{1-\varepsilon^2}}{\varepsilon}\right)/\ln\gamma, \quad \gamma = \frac{\sqrt{\kappa}-1}{1+\sqrt{\kappa}}. \tag{8}$$

If $\tilde{A}$ is nonsymmetric but positive definite, i.e., $(\tilde{A}u, u) \ge \delta(u, u)$ for $\delta > 0$ and $u \ne 0$, then system (3) can be solved by the semiconjugate residual (SCR) method, which is a stable modification of the generalized conjugate residual method described in [5]. In SCR the vectors u^{n+1} and r^{n+1} are calculated by formulas (5), where the coefficients α_n are determined by (6) at $s = 1$ and p^{n+1} are determined by the "long" recursions

$$p^{n+1,0} = r^{n+1}, \quad p^{n+1,l} = p^{n+1,l-1} - \beta_{n,l}p^{l-1},$$
$$l = 1, 2, \ldots, n,$$
$$\beta_{n,l} = -\frac{(\tilde{A}p^l, \tilde{A}p^{n+1,l-1})}{(\tilde{A}p^l, \tilde{A}p^l)}, \quad p^{n+1} = p^{n+1,n}. \tag{9}$$

Institute of Computational Mathematics and Mathematical Geophysics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, pr. Akademika Lavrent'eva 6, 630090 Russia
e-mail: ilin@sscc.ru

doi:10.1016/j.jmb.2007.05.043 J. Mol. Biol. (2007) 371, 873–882

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Nucleoside Analogs 4′*C*-Methyl Thymidine and 4′*C*-Ethyl Thymidine Block DNA Synthesis by Wild-type HIV-1 RT and Excision Proficient NRTI Resistant RT Variants

Paul L. Boyer[1], John G. Julias[2], Zandrea Ambrose[1] Maqbool A. Siddiqui[3], Victor E. Marquez[3] and Stephen H. Hughes[1]*

[1]*HIV Drug Resistance Program National Cancer Institute Frederick, MD 21702-1201 USA*

[2]*SAIC-Frederick, Frederick MD 21702, USA*

[3]*Laboratory of Medicinal Chemistry, Center for Cancer Research, NCI-Frederick, NIH Frederick, MD 21702, USA*

*Corresponding author

HIV-1 can become resistant to nucleoside analogs by developing an enhanced ability to excise the analogs after they have been incorporated. Excision requires that the analog be located at the 3′ terminus of the primer. We have describe nucleoside analogs that do not block DNA synthesis at the point of incorporation, but only after additional dNTPs have been added to the DNA. These nucleoside analogs are called "delayed chain terminators" and are relatively effective inhibitors of drug-resistant HIV-1 reverse transcriptases (RTs) that are excision proficient. However, the first delayed chain terminator that we characterized was poorly phosphorylated in cultured cells. We have examined other nucleoside analogs to determine whether these compounds also act as delayed chain terminators, but were more efficiently converted to the triphosphate form by cellular kinases.

These analogs contain substitutions on the deoxyribose sugar ring at the 4′ carbon (4′C-methyl dT and 4′C-ethyl dT). Unlike true delayed chain terminators, which terminate DNA synthesis in a spatial sense (DNA synthesis is halted only after additional dNTPs have been incorporated after the analog), 4′C-methyl dTTP causes a pause in DNA synthesis at the point of incorporation. However, HIV-1 RT can eventually extend the primer blocked by the 4′ C-Me dTMP analog. 4′C-methyl dTTP blocks DNA synthesis in a temporal sense, rather than in a spatial sense. A primer blocked by 4′C-ethyl dTMP is not extended by HIV-1 RT, and this compound acts like a conventional chain terminator, despite the presence of a 3′-OH group. These compounds effectively block the replication of an HIV-1-based vector that replicates using wild-type HIV-1 RT, but only in the presence of herpes simplex virus thymidine kinase (HSV TK). These compounds are effective against many NRTI drug-resistant RT variants; however, the M184V mutant is relatively resistant.

Published by Elsevier Ltd.

Keywords: HIV-1; reverse transcriptase; nucleoside analogs; drug resistance

Introduction

Although there are a large number of anti-HIV-1 drugs, the virus can develop resistance to all of the drugs. A major class of reverse transcriptase (RT) inhibitors is the nucleoside analogs (NRTIs), which act as chain terminators of viral DNA synthesis when incorporated during reverse transcription. An important mechanism of resistance to NRTIs is that

Abbreviations used: HSV-TK, herpes simplex virus thymidine kinase; NRTIs, nucleoside reverse transcriptase inhibitors; RT, reverse transcriptase.

E-mail address of the corresponding author: hughes@ncifcrf.gov

0022-2836/$ - see front matter. Published by Elsevier Ltd.

Formulas (5) and (9) produce A^tA-orthogonal vectors $p^0, p^1, \ldots, p^{n+1}$ (A^t denotes the transpose matrix) by using modified Gram–Schmidt orthogonalization [6]. In this case, the functional $\Phi_n^{(1)}(r^n) = (r^n, r^n)$ is minimized in $\mathcal{H}_{n+1}(r^0, \tilde{A})$ and the residual vectors r^n are right semiconjugate in the sense that $(\tilde{A} r^k, r^n) = 0$ for $k < n$.

MULTIPLICATIVE PROJECTION METHODS

Denote by $\Omega = \{i = 1, 2, \ldots, N\}$ the set of row indices of A and by Ω_p ($p = 1, 2, \ldots, l$) its disjoint subsets with the numbers of elements m_p such that $\Omega = \bigcup_{p=1}^{l} \Omega_p$ and $m_1 + \ldots + m_l = N$. Accordingly, we introduce the m_p-dimensional subvectors $u_{(p)}, f_{(p)}$ ($p = 1, 2, \ldots, l$) and the rectangular $m_p \times N$ submatrices $A_{(p)}$:

$$u_{(p)} = \{u_i, i \in \Omega_p\}, \quad f_{(p)} = \{f_i, i \in \Omega_p\}, \quad A_{(p)} = \{A_i, i \in \Omega_p\}, \tag{10}$$

where A_i is the ith row of A. Then SLAE (1) can be written as

$$A_{(p)}u = f_{(p)}, \quad p = 1, 2, \ldots, l. \tag{11}$$

To solve it, we consider an iterative process in which the computation of every nth approximation consists of l steps:

$$u^{n,p} = u^{n,p-1} + \omega A_{(p)}^{+} r_{(p)}^{n,p-1}, \quad n = 1, 2, \ldots, \quad p = 1, 2, \ldots, l, \quad u^n = u^{n,l}. \tag{12}$$

Here, $u^{0,0} = \{u_i^0, i = 1, 2, \ldots, N\}$ is the initial vector, ω is an iteration parameter, $r_{(p)}^{n,p-1} = f_{(p)} - A_{(p)}u^{n,p-1}$ is the residual subvector of order m_p, and A is the pseudoinverse of $A_{(p)}$ defined as $A_{(p)}^{+} = A_{(p)}^{t}(A_{(p)}A_{(p)}^{t})^{-1}$ (if $A_{(p)}$ is of full rank m).

It follows that $I - A_{(p)}^{+}A_{(p)}$ is the symmetric positive semidefinite matrix of orthogonal projection onto the pth subspace, which is geometrically represented as the union of the subspaces defined by the rows A_i, $i \in \Omega_p$.

Iterative method (12) can be written in matrix form as

$$u^n = Bu^{n-1} + g, \quad B = (I - T_l)\ldots(I - T_1), \quad T_p = \omega A_{(p)}^{+} A_{(p)}. \tag{13}$$

When $\omega = 1$ and $m_p = 1$, projection method (12), (13) becomes Kaczmarz's "pointwise" algorithm proposed in [7]. It was generalized and studied by different authors (see [8, 9]).

In [10] the following result was proved for an abstract iterative multiplicative projection method of form (13).

Theorem 1. *Let T_p ($p = 1, 2, \ldots, l$) be SPD matrices, and let $v \in \mathbf{R}^N$ be an arbitrary vector such that*

$$\frac{(T_p v, v)}{(v, v)} \le \alpha < 2, \quad p = 1, 2, \ldots, l;$$

$$\|v\| \le \beta \sum_{p=1}^{l} (T_p v, v).$$

Then the 2-norm B in (13) satisfies the estimate

$$\|B\| \le \rho = 1 - \frac{2 - \alpha}{\beta[l + \alpha^2 l(l-1)/2]}.$$

If the matrices $\bar{T}_p = \omega^{-1}T_p$ satisfy the conditions

$$\frac{(\bar{T}_p v, v)}{(v, v)} \le \bar{\alpha} < 2,$$

$$\|v\| \le \bar{\beta}[(\bar{T}_1 v, v) + \ldots + (\bar{T}_l v, v)]$$

for all p, then the choice of $\omega = \dfrac{1}{\bar{\omega}\sqrt{(l-1)l}}$ yields $\rho = 1 - (3\bar{\beta}\bar{\alpha}l)^{-1}$.

Note that, since T_p is not permutable, the transition matrix B of iterative process (13) is not symmetric.

Now consider Kaczmarz's alternating block method, in which every iteration consists of two half-steps. The first is given by formulas (12) or (13), while the second consists of similar computations in reverse order with respect to p:

$$u^{n+1/2,p} = u^{n,p-1} + \omega A_{(p)}^{+} r_{(p)}^{n,p-1}, \quad p = 1, 2, \ldots, l, \quad u^{n+1/2} = u^{n+1/2,l} = u^{n+1/2,l+1}, \quad u^{n+1,p} = u^{n+1/2,p+1} + \omega A_{(p)}^{+} r_{(p)}^{n+1/2,p+1}, \quad p = l, l-1, \ldots, 2, 1, \quad u^{n+1} = u^{n+1,1}. \tag{14}$$

The transition matrix in iterations (14) is the matrix product $B = B_2B_1$, where B_1 denotes B in (13) and B_2 has the form

$$u^{n+1} = B_2B_1u^n + g, \quad B_2 = (I - T_1)(I - T_2)\ldots(I - T_l) = B_1^t. \tag{15}$$

Under the assumptions of Theorem 1, each of the matrices B_1 and B_2 satisfies the estimate $\|B_k\| \le \rho$. Therefore, for the transition matrix of alternating method (14), we have $\|B\| \le \|B_1\| \cdot \|B_2\| \le \rho^2 < 1$.

Since method (14) can be represented in the form of (2) with a SPD matrix B, the convergence of the iterations can be accelerated by formally applying conjugate direction algorithms to preconditioned SLAE (3), which yields the following result.

Theorem 2. *Under the assumptions of Theorem* 1, *the alternating multiplicative projection conjugate gradient methods* (AMPCGM) *defined by* (3), (5), *and* (6) *at* $s = 0, 1$ *and by relations* (14) *and* (15) *converge and* $n(\varepsilon)$ *satisfies estimate* (8), *where* $\kappa = \dfrac{1+\rho^2}{1-\rho^2}$ *and* ρ *is defined in Theorem* 1.

Now consider the sequential projection semiconjugate residual method (SPSCRM), which is an alternative to AMPCGM and is based on the acceleration of algorithm (13) with a nonsymmetric transition matrix B by applying iterations in Krylov subspaces by formulas (5) and (9), where the preconditioned matrix is determined by (3) and (13). A specific feature of SPSCRM is that, like in the GMRES algorithm [4], all the previous vectors $p^0, p^1, \ldots, p^n$ have to be stored to determine u^{n+1}. These two methods have identical convergence properties, since they both ensure the minimality of $\Phi_n^{(1)}(r^n)$ in the subspace $\mathcal{H}_{n+1}(r^0, \tilde{A})$. The following result holds for the given multiplicative method.

Theorem 3. *Suppose that the multiplicative SPSCRM algorithm defined by formulas* (3), (5), (6) *at* $s = 1$ *and by* (9) *and* (11)–(13) *has a diagonalizable matrix* $\tilde{A} = X\Lambda X^{-1}$, $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$, *where* λ_i *are the eigenvalues of* $\tilde{A}$ *and* X *is a square matrix whose columns are the corresponding eigenvectors. Assume that the conditions of Theorem* 1 *hold.*

Then this method converges and $n(\varepsilon)$ *satisfies the estimate*

$$n(\varepsilon) \le 1 + \left(\ln \frac{1+\sqrt{1-\varepsilon_1^2}}{\varepsilon_1} \right) / \ln \frac{\gamma_1}{\gamma_2},$$

$$\varepsilon_1 = \frac{\varepsilon}{(\|X\|_2 \cdot \|X^{-1}\|_2)}.$$

Here, $\gamma_1 = a + \sqrt{a^2 - d^2}$; $\gamma_2 = c + \sqrt{c^2 - d^2}$; *and* c, d, *and* a *are the coordinate of the center, the focal distance* ($d^2 < c^2$), *and the major semiaxis of an ellipse in the complex plane that contains all* λ_i.

Note that, like for GMRES, modifications with a constrained number of stored direction vectors can be designed for SPSCRM. This reduces the computational costs per iteration but increases $n(\varepsilon)$.

ADDITIVE PROJECTION METHODS

Versions of Cimmino's iterative method were considered in [11–13] (see also the references therein). Its elementary step, as in Kaczmarz's algorithm, consists of projecting a point of an N-dimensional space onto the hyperplane described by the ith equation of the original SLAE. However, these operations are sequential in the former case while being simultaneous in the latter: for a given approximation u^n, all the projections $u^{n,i}$ onto the hyperplane A_i are found, and the new u^{n+1} is determined as their linear combination. This additive projection algorithm in its block version for solving SLAE (11) is written as

$$u^{n,p} = u^n + A_{(p)}^+ r_{(p)}^{n-1}, \quad p = 1, 2, \ldots, l,$$
$$u^{n+1} = \frac{u^{n,1} + u^{n,2} + \ldots + u^{n,l}}{l}, \tag{16}$$

which is equivalent to the matrix form

$$u^{n+1} = Bu^n + g, \quad B = I - l^{-1} \sum_{p=1}^{l} A_{(p)}^+ A_{(p)}$$
$$= I - l^{-1} \sum_{p=1}^{l} T_p, \quad g = l^{-1} \sum_{p=1}^{l} A_{(p)}^+ f_{(p)}. \tag{17}$$

Obviously, the limiting vector of this iterative process $u = \lim_{n \to 0} u^n$ satisfies the preconditioned system of equations

$$\tilde{A}u = \tilde{f}, \quad \tilde{A} = \sum_{p=1}^{l} T_p, \quad \tilde{f} = \sum_{p=1}^{l} A_{(p)}^+ f_{(p)}. \tag{18}$$

The eigenvalues of $\tilde{A}$ are estimated as follows (see [10]).

Theorem 4. *Let* $0 < \alpha < 2$ *and* $0 < \rho < 1$ *be the same as in Theorem* 1.

Then the eigenvalues $\lambda(\tilde{A})$ *of the matrix* $\tilde{A}$ *in* (18) *satisfy* $\dfrac{(2-\alpha)(1-\rho)}{4} \le \lambda(\tilde{A}) \le \alpha l$.

This assertion implies the following result.

Theorem 5. *For conjugate gradient methods* (5), (6) *as applied to SLAE* (18) *of the additive projection algorithm, estimate* (8) *for* $n(\varepsilon)$ *holds true, in which the condition number* $\kappa(\tilde{A})$ *satisfies the inequality* $\kappa \le \dfrac{4\alpha l}{(2-\alpha)(1-\rho)}$.

Obviously, multiplication by $\tilde{A}$ in this algorithm corresponds to one iteration in Cimmino's block method (17).

Remark 1. Theorems 2 and 5 imply that the additive projection method converges more slowly that the multiplicative one. However, additive algorithms are more advantageous in parallelization, since the projections $u^{n,p}$ can be computed simultaneously on different processors.

Remark 2. Theorems 1 and 4 above were used in [10] to substantiate multiplicative and additive domain decomposition methods. Obviously, Theorems 2, 3, and 5 on conjugate and semiconjugate direction methods with acceleration also hold for these applications.

A promising direction in the acceleration of iterations, for example, in the domain decomposition method (see [2]) is the use of algorithms in Krylov subspaces with dynamic preconditioning. A generalization of the methods discussed above is the nonstationary iterative process

$$u^{n+1} = B_u u^n + g^n = u^n + C_n^{-1}(f - Au^n), \quad B_n = I - C_n^{-1}A, \tag{19}$$

where C_n are easily invertible preconditioning matrices. The acceleration of the corresponding iterations in the subspaces $\mathcal{K}_{n+1}(r^0, C_n^{-1}A) = \operatorname{span}\{C_0^{-1}r^0, AC_1^{-1}r^0, \ldots, A^nC_n^{-1}r^0\}$ is ensured by the dynamically preconditioned semiconjugate direction (DPSCD) method

$$r^0 = f - Au^0, \quad p^0 = C_0^{-1}r^0, \quad n = 0, 1, \ldots;$$
$$u^{n+1} = u^n + \alpha_n p^n, \quad r^{n+1} = r^n - \alpha_n Ap^n,$$
$$p^{n+1} = C_{n+1}^{-1}r^{n+1} + \sum_{k=0}^{n-1}\beta_{n,k}p^n = p^{n+1,l} + \sum_{k=l}^{n-1}\beta_{n,k}p^n, \tag{20}$$
$$\alpha_n = \frac{(Ar^n, r^n)}{(Ap^n, Ap^n)}, \quad \beta_{n,k} = -\frac{(Ap^k, Ap^{n,k})}{(Ap^k, Ap^k)}.$$

The DPSCD algorithm minimizes the residual norm $\|r^{n+1}\|$ in the subspace $\mathcal{K}_{n+1}(r^0, C_n^{-1}A)$ so that

$$\|r^{n+1}\|^2 = (r^0, r^0) - \frac{(AC_0^{-1}r^0, r^0)^2}{(Ap^0, Ap^0)} - \ldots - \frac{(AC_n^{-1}r^n, r^n)^2}{(Ap^n, Ap^n)}.$$

Thus, if the matrices $C_n^{-1}A$ are positive definite, the DPSCD method converges. However, additional studies are required to estimate $n(\varepsilon)$ and optimize the iterations. Note that the matrices B_n in (19) can be defined, for example, as the products of the arbitrary factors $(I - T_k)$ in (13), but the different B_n must include all the rows A_i of the original matrix A.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 05-01-10487) and by the Department of Mathematical Sciences of the Russian Academy of Sciences (program no. 1.3.5).

REFERENCES

1. G. Golub and C. Van Loan, *Matrix Computations* (Johns Hopkins Univ. Press, Baltimore, 1989).
2. O. Axelsson, *Iterative Solution Methods* (Cambridge Univ. Press, New York, 1994).
3. V. P. Il'in, *Incomplete Factorization Methods for Algebraic Systems* (Nauka, Moscow, 1995) [in Russian].
4. Y. Saad, *Iterative Methods for Sparse Linear Systems* (PWS, new York, 1996).
5. S. C. Eisenstat, H. C. Elman, and M. H. Schultz, SIAM J. Numer. Anal. **20**, 345–357 (1983).
6. V. P. Il'in, *Numerical Analysis* (Inst. Vychisl. Mat I Mat. Geofiz. Sib. Otd. Ross. Akad. Nauk, Novosibirsk, 2004), part 1 [in Russian].
7. S. Kaczmarz, Bull. Int. Acad. Polon. Sci. Lett. Gl. Sci. Math. Nat. A, 355–357 (1937).
8. K. Tanabe, Numer. Math. **17**, 203–214 (1971).
9. V. P. Il'in, Sib. Zh. Ind. Mat. **9** (3), 39–49 (2006).
10. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, Math. Comput. **57** (195), 1–21 (1991).
11. C. Cimmino, Ric. Sci. Progr. Tech. Econ. Naz. **16**, 326–333 (1938).
12. R. Bramley and A. Sameh, SIAM J. Sci. Stat. Comput. **13**, 168–193 (1992).
13. G. Appleby and D. C. Smolarski, Electron. Trans. Numer. Anal. **20**, 243–275 (2005).

SPELL: OK