Direct Inversion of the Iterative Subspace

The Molecular Sciences Software Institute

Daniel G. A. Smith

I.   DIRECT INVERSION OF THE ITERATIVE SUBSPACE CONVERGENCE ACCELERATION

Direct Inversion of the Iterative Subspace (DIIS) is often used for convergence acceleration[?] of SCF wavefunctions. For iterative diagonalization, we discard all previous guess vectors (orbitals) and generate completely new guess vectors at each iteration. DIIS keeps previous guess vectors and builds the best guess as a linear combination of all previous guess vectors.

For the explanation of DIIS, we will break with all conventional terminology to stress that this is a linear algebra technique that is widely applicable to many problems besides SCF, as we will see in the next several sections. Let us assume that we have generated several state vectors (s) with a series of complementary error vectors (r) through some iterative method.

For DIIS to be applicable, we assume that both the next residual and the next state vector can be built as a linear combination of previous guess vectors with coefficients c.

$$s_{i+1} = \sum_i^n c_i s_i \tag{1}$$

$$r_{i+1} = \sum_i^n c_i r_i \tag{2}$$

$$\sum_i^n c_i = 1 \tag{3}$$

Therefore, we want to select the coefficients in a manner that minimizes the norm of the residual

$$r_{i+1} \cdot r_{i+1} = \sum_{ij}^n c_i c_j (r_i \cdot r_j) \tag{4}$$

This can be solved through the following matrix equations

$$\begin{bmatrix} B_{11} & B_{12} & ... & B_{1n} & -1 \\ B_{21} & B_{22} & ... & B_{2n} & -1 \\ ... & ... & ... & ... & ... \\ B_{n1} & B_{n2} & ... & B_{nn} & -1 \\ -1 & -1 & ... & -1 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ ... \\ c_n \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ ... \\ 0 \\ -1 \end{bmatrix} \tag{5}$$

where B is a matrix of inner products between error vectors

$$B_{ij} = (r_i \cdot r_j) \tag{6}$$

and λ is value of the Lagrange multiplier. It should be noted that often our error vectors are in effect tensors, e.g., $F_{\mu\nu}$, and the norm is computed as the generalized inner product between tensors.

For SCF, the state vector is always the Fock matrix and the error vector is typically the orthonormalized orbital gradient expressed in a AO matrix formalism,

$$r_{\mu\nu} = (\mathbf{A}^T(\mathbf{FDS} - \mathbf{SDF})\mathbf{A})_{\mu\nu} \tag{7}$$

The procedure for applying DIIS to SCF at every iteration is therefore:

1. Compute the Fock (s_n) matrix using $\mathbf{C}_n$ orbitals

2. Construct the AO gradient (r_n)

3. Obtain the DIIS Fock guess (s_{DIIS})

4. Obtain $\mathbf{C}_{n+1}$ orbitals by diagonalizing the DIIS Fock guess (s_{DIIS})

This procedure is iterated until the desired convergence thresholds are reached.

FIG. 1. A RHF computation of "physicist's water molecule" starting with a core Hamiltonian guess in the cc-pVDZ basis set utilizing DIIS convergence.

```
RHF Iteration    1: Energy = -68.98003273414295    dE = -6.898E+01    dRMS = 1.165E-01
RHF Iteration    2: Energy = -69.64725442845806    dE = -6.672E-01    dRMS = 1.074E-01 DIIS
RHF Iteration    3: Energy = -75.79192914624532    dE = -6.144E+00    dRMS = 2.892E-02 DIIS
RHF Iteration    4: Energy = -75.97218922804181    dE = -1.802E-01    dRMS = 7.564E-03 DIIS
RHF Iteration    5: Energy = -75.98936905846086    dE = -1.717E-02    dRMS = 8.749E-04 DIIS
RHF Iteration    6: Energy = -75.98971633493079    dE = -3.472E-04    dRMS = 5.356E-04 DIIS
RHF Iteration    7: Energy = -75.98979323982247    dE = -7.690E-05    dRMS = 6.212E-05 DIIS
RHF Iteration    8: Energy = -75.98979567508871    dE = -2.435E-06    dRMS = 1.972E-05 DIIS
RHF Iteration    9: Energy = -75.98979578301157    dE = -1.079E-07    dRMS = 1.727E-06 DIIS
```

The DIIS convergence procedure is then applied to the "physicist's water molecule" in Fig. 1. It should be noted that while DIIS can be utilized on every iteration, the results are equal to pure diagonalization until two error vectors have been constructed. Therefore, the first iteration which benefits from DIIS acceleration is iteration three. As can be seen, roughly the same energy convergence is obtained in 9 iterations compared to 24 iterations without DIIS. The cost of DIIS itself is negligible; however, the cost of building the gradient is on the order of N^3, which is still much less than the Coulomb and Exchange matrix builds.

REFERENCES

[1] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).