General Contraction Basis Sets: fast direct Fock Matrix formation

Alexander A. Granovsky

Laboratory of Chemical Cybernetics, M.V. Lomonosov Moscow State University, Moscow, Russia

May 04, 2005

Fock Matrix

- $F_2(D) = J(D) - K(D)$
  - $J_{pq} = \Sigma_{rs}(pq|rs)D_{rs}$
  - $K_{pr} = \Sigma_{qs}(pq|rs)D_{qs}$

    where (pq|rs) are 2-e integrals in AO basis
- Note: required not only for HF/DFT
  - CIS, TDHF, TDDFT, CI, MCSCF, ...
  - Orbital response equations for properties
  - Molecular gradients: orbital response equations for all non fully orbital-variational methods
  - Molecular hessians: orbital response equations for all methods

GTO Basis Sets

- Segmented Contraction (SC)
  - $\chi_i = \Sigma_j\, c_{ij} x^k y^l z^m \exp(-\alpha_{ij} r^2)$
  - Examples: 6-31*, etc...
- General Contraction (GC)
  - $\chi_i = \Sigma_j\, c_{ij} x^k y^l z^m \exp(-\alpha_j r^2)$
    - **common set of exponents within each shell type**
    - **(very) long expansions**
  - Examples: Roos’ ANO, Dunning’s cc-pVXZ (to some degree)

Standard Direct Fock matrix formation code

```
loop over all shells (P index)
    loop over Q ≤ P, Q in P domain
        loop over all shells (R index)
            loop over S ≤ R, S in R domain
                Check (PQ|RS) using selection criterion
                Calculate non-negligible (PQ|RS) if any
                Add their contributions to the Fock matrix
            end loop
        end loop
    end loop
end loop
```

Simplified code calculating four-center integrals over shell quartet

```
loop over all primitives in P shell
    loop over all primitives in Q shell
        loop over all primitives in R shell
            loop over all primitives in S shell
                Calculate contribution due to combination of four
                primitive functions to (PQ|RS)
            end loop
        end loop
    end loop
end loop
```

Preliminary considerations

- Standard code is very bad for GC basis sets:
  - does not reuse already calculated integrals over common primitives
  - does not like long contractions, especially for high L.
  - does not use integral screening efficiently

Modified Fock matrix formation code, most common way

```
loop over all shell subsets (P index)
    loop over Q subsets ≤ P, Q in P domain
        loop over all shell subsets (R index)
            loop over S subsets ≤ R, S in R domain
                Calculate all non-negligible (pq|rs) combinations
                Perform small four-index transformation to get all (PQ|RS)
                Save on disk or add contributions to the Fock matrix
            end loop
        end loop
    end loop
end loop
```

Common way problems:

- Requires non-trivial extra logic in loops
- Complex algorithms
- Non-trivial implementation of:
  - parallel mode algorithms
  - integral screening
  - use of molecular symmetry
- Can be a memory hog
  - Example: g-functions, 10 primitives:
    - $15^4 * 10^4 \approx$ 500 000 000 words of memory are required only to store all intermediate (pq|rs)

Alternative way to direct Fock Matrix construction:

- We do not need 2-e integrals themselves, only their contributions to the Fock matrix
- Let us consider transformation from (pq|rs) to (PQ|RS) as four-index transformation using specific transformation matrix:
  - $(PQ|RS) = \Sigma_p \Sigma_q \Sigma_r \Sigma_s C_{Pp}C_{Qq}C_{Rr}C_{Ss}(pq|rs)$
    - Matrix C is completely defined by the basis set contraction coefficients
    - Matrix C is block-diagonal (this will not be very important)

Alternative way to direct Fock Matrix construction, contd.:

- $(PQ|RS) = \Sigma_p \Sigma_q \Sigma_r \Sigma_s C_{Pp}C_{Qq}C_{Rr}C_{Ss}(pq|rs)$
- $J_{PQ} = \Sigma_{RS}(PQ|RS)D_{RS}$
- $J_{PQ} = \Sigma_{RS} \Sigma_p \Sigma_q \Sigma_r \Sigma_s C_{Pp}C_{Qq}C_{Rr}C_{Ss}(pq|rs)\ D_{RS}$
- $J_{PQ} = \Sigma_p \Sigma_q C_{Pp}C_{Qq}\Sigma_r \Sigma_s (pq|rs)\ \Sigma_{RS} C_{Rr}C_{Ss}D_{RS}$
- Let us define:
  - $\mathit{D}_{rs}=\Sigma_{RS} C_{Rr}C_{Ss}D_{RS}$ <=> $\mathit{D} = C^{+}DC$
  - $\mathit{J}_{pq} = \Sigma_r \Sigma_s (pq|rs)\ \mathit{D}_{rs}$
- Then:
  - $J_{PQ} = \Sigma_p \Sigma_q C_{Pp}C_{Qq}\ \mathit{J}_{pq}$ <=> $J = CJC^{+}$

How it works:

- Having initial GC basis set and density matrix:
  - Define intermediate (fully or partially uncontracted) basis
  - Define constant transformation matrix C
  - Transform density to the intermediate basis set
  - Calculate *F* and/or *J* & *K* matrix (matrices) using standard direct SCF code
- Finally, transform *F* (*J*, *K*) back to the initial basis set

Example: FeO molecule/Roos Augmented Triple Zeta ANO

- Fe:
  - (21s,15p,10d,6f,4g) -> [8s,7p,5d,3f,2g]
- O:
  - (14s,9p,4d,3f) -> [5s,4p,3d,2f]
- FeO:
  - 39 shells, 174 Cartesian/139 spherical basis functions

Example, contd.

- PC GAMESS direct RHF calculations, single-CPU Opteron 144
  - Using GAMESS (US) 2-e code:
    - First iteration: 175 min., last iteration: 175 min.
  - Using PC GAMESS 2-e code:
    - First iteration: 83 min., last iteration: 83 min.
  - <u>Using PC GAMESS GENCON code:</u>
    - <u>First iteration: 51 sec., last iteration: 33 sec.</u>

Thank you for your attention!