

# CONJUGATE GRADIENT FORMULAS

**Neculai Andrei,**  
*Research Institute for Informatics,  
 Center for Advanced Modeling and Optimization,  
 8-10, Averescu Avenue, Bucharest 1.  
 E-mail: nandrei @ ici.ro*

Nr.	File	Formula	Name
1.	Z1	$\beta_k^{HS} = \frac{\mathbf{y}_k^T \mathbf{g}_{k+1}}{\mathbf{y}_k^T \mathbf{s}_k}$	Hestenes and Stiefel (HS)
2.	Z2	$\beta_k^{FR} = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k}$	Fletcher and Reeves (FR)
3.	Z3	$\beta_k^{PRP} = \frac{\mathbf{y}_k^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k}$	Polak-Ribiere and Polyak (PRP)
4.	Z4	$\beta_k^{PRP+} = \max\left\{0, \frac{\mathbf{y}_k^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k}\right\}$	Polak-Ribiere and Polyak + (PRP+)
5.	Z5	$\beta_k^{CD} = -\frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{d}_k}$	Conjugate Descent – Fletcher (CD)
6.	Z6	$\beta_k^{LS} = -\frac{\mathbf{y}_k^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{d}_k}$	Lui and Storey (LS)
7.	Z7	$\beta_k^{DY} = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1}}{\mathbf{y}_k^T \mathbf{s}_k}$	Dai and Yuan (DY)
8.	Z8	$\beta_k^{DL} = \frac{\mathbf{g}_{k+1}^T (\mathbf{y}_k - t \mathbf{s}_k)}{\mathbf{y}_k^T \mathbf{s}_k}$	Dai and Liao (DL)
9.	Z9	$\beta_k^{DL+} = \max\left\{0, \frac{\mathbf{y}_k^T \mathbf{g}_{k+1}}{\mathbf{y}_k^T \mathbf{s}_k}\right\} - t \frac{\mathbf{s}_k^T \mathbf{g}_{k+1}}{\mathbf{y}_k^T \mathbf{s}_k}$	Dai and Liao + (DL+)
10.	Z10	$\beta_k^{ASDC} = \frac{\mathbf{g}_{k+1}^T \mathbf{g}_{k+1} - (\mathbf{y}_k^T \mathbf{g}_{k+1})(\mathbf{s}_k^T \mathbf{g}_{k+1})}{\mathbf{y}_k^T \mathbf{s}_k (\mathbf{y}_k^T \mathbf{s}_k)^2}$	Andrei Sufficient Descent Condition. Please, see the paper for AML, AML5382. ( $\mathbf{y}_k^T \bar{\mathbf{Q}}_{k+1} = 0$ )
11.	Z11	$\beta_k^{hDY} = \max\left\{c \beta_k^{DY}, \min\left\{\beta_k^{HS}, \beta_k^{DY}\right\}\right\}$	Hybrid DY (hDY)
12.	Z12	$\beta_k^{hDYz} = \max\left\{0, \min\left\{\beta_k^{HS}, \beta_k^{DY}\right\}\right\}$	Hybrid DY zero (hDYz)
13.	Z13	$\beta_k^{GN} = \max\left\{-\beta_k^{FR}, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}$	Gilbert and Nocedal (GN)
14.	Z14	$\beta_k^{HuS} = \max\left\{0, \min\left\{\beta_k^{PRP}, \beta_k^{FR}\right\}\right\}$	Hu and Storey (HuS)

15.	Z15	$\beta_k^{TaS} = \begin{cases} \beta_k^{PRP} & 0 \leq \beta_k^{PRP} \leq \beta_k^{FR}, \\ \beta_k^{FR} & \text{otherwise} \end{cases}$	Touati-Ahmed and Storey (TaS)
16.	Z16	$\beta_k^{LS-CD} = \max\{0, \min\{\beta_k^{LS}, \beta_k^{CD}\}\}$	Hybrid LS, CD (LS-CD)
17.	Z17	$\beta_k^{BM} = \frac{(\theta y_k - s_k)^T g_{k+1}}{y_k^T s_k}$	Birgin and Martinez, Scaled Perry
18.	Z18	$\beta_k^{BM+} = \max\left\{0, \frac{\theta y_k^T g_{k+1}}{y_k^T s_k}\right\} - \frac{s_k^T g_{k+1}}{y_k^T s_k}$	Brigin and Martinez +
19.	Z19	$\beta_k^{sPRP} = \frac{\theta_{k+1} y_k^T g_{k+1}}{\alpha_k \theta_k g_k^T g_k}$	Scaled Polak-Ribiere-Polyak (sPRP)
20.	Z20	$\beta_k^{sFR} = \frac{\theta_{k+1} g_{k+1}^T g_{k+1}}{\alpha_k \theta_k g_k^T g_k}$	Scaled Fletcher-Reeves (sFR)
21.	Z21	$\beta_k^A = \frac{1}{y_k^T s_k} \left( y_k^T g_{k+1} - \frac{(y_k^T y_k)(s_k^T g_{k+1})}{g_k^T g_k} \right)$	Andrei ( <i>Sufficient descent condition from PRP</i> ) Please, see Remark 8.3.3 and formula (8.3.130) in the book: “ <i>Criticism of the Unconstrained Optimization Algorithms Reasoning</i> ”.
22.	Z22	$\beta_k^{ACGSD} = \frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(y_k^T g_{k+1})(s_k^T g_{k+1})}{(y_k^T s_k)^2}$	Andrei ( <i>Sufficient descent condition from DY</i> ) (ACGSD) Please see paper for SIOPT, #067836. ( $y_k^T \bar{Q}_{k+1} g_{k+1} = 0$ )
23.	Z23	$\beta_k^{ACGSDz} = \max\left\{0, \frac{y_k^T g_{k+1}}{y_k^T s_k}\right\} \left(1 - \frac{s_k^T g_{k+1}}{y_k^T s_k}\right)$	Andrei ( <i>Sufficient descent condition from DY zero</i> ) (ACGSDz)

### Remarks:

1. All these conjugate gradient algorithms are implemented in CGALL.FOR package.
2. The formula  $\beta_k^{ASDC} = \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} - \frac{(y_k^T g_{k+1})(s_k^T g_{k+1})}{(y_k^T s_k)^2}$  is obtained from the conjugacy condition

$$y_k^T d_{k+1} = 0, \text{ where}$$

$$d_{k+1} = -Q_{k+1} g_{k+1},$$

and

$$Q_{k+1} = \theta_{k+1} I - \frac{s_k g_{k+1}^T}{y_k^T s_k} + \delta_k \frac{\|g_{k+1}\|^2}{(y_k^T s_k)^2} (s_k s_k^T).$$

Now, by symmetrization of  $Q_{k+1}$  as

$$\bar{Q}_{k+1} = \theta_{k+1} I - \frac{s_k g_{k+1}^T + g_{k+1} s_k^T}{y_k^T s_k} + \delta_k \frac{\|g_{k+1}\|^2}{(y_k^T s_k)^2} (s_k s_k^T)$$

and considering the conjugacy condition  $y_k^T d_{k+1} = 0$ , i.e.  $y_k^T \bar{Q}_{k+1} = 0$ , after some algebra we get:

$$\theta_{k+1} = \frac{g_{k+1}^T g_{k+1}}{y_k^T g_{k+1}} \quad \text{and} \quad \delta_k = \frac{y_k^T g_{k+1}}{g_{k+1}^T g_{k+1}} = \frac{1}{\theta_{k+1}}$$

with which we get  $\beta_k^{ASDC}$  as above.

Please see the paper:

**Neculai Andrei**, “*Dai-Yuan conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization*”, Submitted AML, November 20, 2006. (Submitted)

3. Formula  $\beta_k^A = \frac{1}{y_k^T s_k} \left( y_k^T g_{k+1} - \frac{(y_k^T y_k)(s_k^T g_{k+1})}{g_k^T g_k} \right)$  is obtained by a modification of PRP conjugate gradient formula.

4. Formula  $\beta_k^{ACGSD} = \frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(y_k^T g_{k+1})(s_k^T g_{k+1})}{(y_k^T s_k)^2}$  is obtained by a modification of DY conjugate gradient formula. Observe that the direction:

$$d_{k+1} = -g_{k+1} + \frac{g_{k+1}^T g_{k+1}}{y_k^T s_k} s_k - \delta_k \frac{\|g_{k+1}\|^2}{(y_k^T s_k)^2} (g_{k+1}^T s_k) s_k$$

which can be written as

$$d_{k+1} = -Q_{k+1} g_{k+1},$$

where the matrix  $Q_{k+1}$  is:

$$Q_{k+1} = I - \frac{s_k g_{k+1}^T}{y_k^T s_k} + \delta_k \frac{\|g_{k+1}\|^2}{(y_k^T s_k)^2} (s_k s_k^T).$$

Now, by symmetrization of  $Q_{k+1}$  as:

$$\bar{Q}_{k+1} = I - \frac{s_k g_{k+1}^T + g_{k+1} s_k^T}{y_k^T s_k} + \delta_k \frac{\|g_{k+1}\|^2}{(y_k^T s_k)^2} (s_k s_k^T),$$

we can consider the direction

$$d_{k+1} = -\bar{Q}_{k+1} g_{k+1}.$$

From the conjugacy condition,  $y_k^T d_{k+1} = 0$ , i.e.

$$y_k^T \bar{Q}_{k+1} g_{k+1} = 0,$$

after some algebra it follows that

$$\delta_k = \frac{y_k^T s_k}{g_{k+1}^T s_k} + \frac{g_{k+1}^T y_k}{\|g_{k+1}\|^2} - \frac{(g_{k+1}^T y_k)(y_k^T s_k)}{\|g_{k+1}\|^2 (g_{k+1}^T s_k)}.$$

Therefore, we get

$$\beta_k^{ACGSD} = \frac{1}{y_k^T s_k} \left( y_k - \frac{g_{k+1}^T y_k}{y_k^T s_k} s_k \right)^T g_{k+1}.$$

Please see the paper:

**Neculai Andrei**, “*Another nonlinear conjugate gradient algorithm with conjugacy and sufficient descent conditions for unconstrained optimization*”, Submitted SIOPT, December 22, 2006.

February 28, 2007