Quadratic programming (QP)

Oscar S. Dalmau Cedeño
dalmau@cimat.mx

10 de octubre de 2017

QP Problem

The quadratic programming problem is formulates as follows

$$
\begin{aligned}
\text{mín}\, q(\boldsymbol{x}) &= \frac{1}{2}\boldsymbol{x}^T\mathbf{G}\boldsymbol{x} - \boldsymbol{c}^T\boldsymbol{x} && (1)\\
s.a.:\ \boldsymbol{a}_i^T\boldsymbol{x} &= b_i,\ i \in \mathcal{E} && (2)\\
\boldsymbol{a}_i^T\boldsymbol{x} &\geq b_i,\ i \in \mathcal{I} && (3)
\end{aligned}
$$

where $\mathbf{G} \in \mathbb{R}^{n\times n}$ is semidefinitive positive matrix (convex case) and $\boldsymbol{c}, \boldsymbol{x}, \{\boldsymbol{a}_i,\ i \in \mathcal{E} \cup \mathcal{I}\} \in \mathbb{R}^n$, $\{b_i,\ i \in \mathcal{E} \cup \mathcal{I}\} \in \mathbb{R}$.

QP KKTs

Active set:

$$\mathcal{A}(\boldsymbol{x}^*) = \{i \in \mathcal{E} \cup \mathcal{I} |\ \boldsymbol{a}_i^T \boldsymbol{x}^* = b_i\} \tag{4}$$

The KKTs conditions at $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ for QP are:

$$\nabla \mathcal{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \mathbf{G}\boldsymbol{x}^* - \boldsymbol{c} - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i^* \boldsymbol{a}_i = 0 \tag{5}$$

$$\boldsymbol{a}_i^T \boldsymbol{x}^* = \boldsymbol{b}_i,\ i \in \mathcal{E} \tag{6}$$

$$\boldsymbol{a}_i^T \boldsymbol{x}^* \geq \boldsymbol{b}_i,\ i \in \mathcal{I} \tag{7}$$

$$\lambda_i^* \geq 0,\ i \in \mathcal{I} \tag{8}$$

$$\lambda_i^* (\boldsymbol{a}_i^T \boldsymbol{x}^* - \boldsymbol{b}_i) = 0, i \in \mathcal{E} \cup \mathcal{I} \tag{9}$$

The QP problem with equality constraints can be written as follows

$$
\begin{aligned}
\min q(\boldsymbol{x}) &= \frac{1}{2}\boldsymbol{x}^T \mathbf{G}\boldsymbol{x} - \boldsymbol{c}^T\boldsymbol{x} && (10)\\
s.a.:\ \mathbf{A}\boldsymbol{x} &= \boldsymbol{b} && (11)
\end{aligned}
$$

where $\mathbf{G} \in \mathbb{R}^{n\times n}$ is semidefinitive positive matrix and $A \in \mathbb{R}^{m\times n}$ is a complete rank matrix.

KKTs

Lagrangian

$$\mathcal{L}(\boldsymbol{x}, \lambda) = \frac{1}{2}\boldsymbol{x}^T\mathbf{G}\boldsymbol{x} - \boldsymbol{c}^T - \lambda^T(\mathbf{A}\boldsymbol{x} - \boldsymbol{b}) \quad (12)$$

KKTs

$$\nabla_x\mathcal{L}(\boldsymbol{x}, \lambda) = \mathbf{G}\boldsymbol{x}^* - \boldsymbol{c} - \mathbf{A}^T\lambda^* = 0 \quad (13)$$

$$\mathbf{A}\boldsymbol{x}^* = \boldsymbol{b} \quad (14)$$

KKTs

KKTs

$$\begin{bmatrix} \mathbf{G} & -\mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^* \\ \lambda^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{b} \end{bmatrix} \tag{15}$$

Karush-Kuhn-Tucker system

Let $\boldsymbol{x}$ an estimate of the solution and $\boldsymbol{p}$ is the desired step then $\boldsymbol{x}^* = \boldsymbol{x} + \boldsymbol{p}$

$$\begin{bmatrix} \mathbf{G} & -\mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} + \boldsymbol{p} \\ \lambda \end{bmatrix} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{b} \end{bmatrix} \qquad (16)$$

then

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{p} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} \qquad (17)$$

with $\mathbf{g} = \mathbf{G}\boldsymbol{x} - \boldsymbol{c}$, $\mathbf{h} = \mathbf{A}\boldsymbol{x} - \boldsymbol{b}$ and $\boldsymbol{x}^* = \boldsymbol{x} + \boldsymbol{p}$. The matrix of the previous problem is called the *Karush-Kuhn-Tucker (KKT) matrix*.

Karush-Kuhn-Tucker matrix

Lemma

Let $\mathbf{A}$ have full row rank, and assume that the reduced Hessian matrix $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ is positive definite. Then the KKT matrix

$$\mathbf{K} = \begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \tag{18}$$

is nonsingular, and hence there is a unique vector pair $(\boldsymbol{x}^*, \lambda^*)$ satisfying the KKT conditions.

Here $\mathbf{Z} \in \mathbb{R}^{n\times(n-m)}$ is a matrix whose columns are a basis for the null space of $\mathbf{A}$. That is, $\mathbf{Z}$ has full rank and satisfies $\mathbf{A}\mathbf{Z} = 0$.

Karush-Kuhn-Tucker matrix

Let $[w; v]$ be a vector of the null space of $\mathbf{K}$.

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{19}$$

$$\mathbf{G}w + \mathbf{A}^T v = 0 \tag{20}$$

$$\mathbf{A}w = 0 \tag{21}$$

From $\mathbf{A}w = 0$ we obtain that w belongs to the null space of $\mathbf{A}$, ie, we can write $w = \mathbf{Z}\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{R}^{n-m}$

Karush-Kuhn-Tucker matrix

Then, using $\mathbf{AZ} = 0$ and $w = \mathbf{Z}\boldsymbol{u}$

$$
\begin{aligned}
\mathbf{Z}^T\mathbf{G}w + \mathbf{Z}^T\mathbf{A}^T v &= 0 \qquad (22)\\
\mathbf{Z}^T\mathbf{G}\mathbf{Z}\boldsymbol{u} &= 0 \qquad (23)
\end{aligned}
$$

as $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ has inverse then $\boldsymbol{u} = 0$ then $w = 0$ and therefore $v = 0$ from $\mathbf{A}^T v = 0$ and the fact that $\mathbf{A}$ is a full rank matrix. Then the null space of $\mathbf{K}$ is the null vector, this implies that $\mathbf{K}$ is invertible.

Theorem

Let $\mathbf{A}$ have full row rank and assume that the reduced-Hessian matrix $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ is positive definite. Then the vector x^* satisfying the KKT conditions is the unique global solution of the optimization problem.

$\mathbf{Z} \in \mathbb{R}^{n\times(n-m)}$ is a matrix whose columns are a basis for the null space of $\mathbf{A}$.

Let $\boldsymbol{x}$ any feasible point and $\boldsymbol{x}^*$ the optimum, we are going to prove that $q(\boldsymbol{x}) \geq q(\boldsymbol{x}^*)$.
Let's define $\boldsymbol{p} = \boldsymbol{x}^* - \boldsymbol{x}$ then $\boldsymbol{b} = \mathbf{A}\boldsymbol{x}^* = \mathbf{A}\boldsymbol{x}$ and $0 = \mathbf{A}(\boldsymbol{x}^* - \boldsymbol{x}) = \mathbf{A}\boldsymbol{p}$ that is $\boldsymbol{p}$ belongs the he null space of $\mathbf{A}$ and we can write $\boldsymbol{p} = \mathbf{Z}\boldsymbol{u}$ for some $\boldsymbol{u}$.
On the other hand, and using the KKT $-\mathbf{G}\boldsymbol{x}^* + \boldsymbol{c} = \mathbf{A}^T\lambda^*$

$$
\begin{aligned}
q(\boldsymbol{x}) &= q(\boldsymbol{x}^* - \boldsymbol{p}) = \frac{1}{2}(\boldsymbol{x}^* - \boldsymbol{p})^T\mathbf{G}(\boldsymbol{x}^* - \boldsymbol{p}) - \boldsymbol{c}^T(\boldsymbol{x}^* - \boldsymbol{p}) \\
&= q(\boldsymbol{x}^*) + \frac{1}{2}\boldsymbol{p}^T\mathbf{G}\boldsymbol{p} + \boldsymbol{p}^T(-\mathbf{G}\boldsymbol{x}^* + \boldsymbol{c}) \\
&= q(\boldsymbol{x}^*) + \frac{1}{2}\boldsymbol{p}^T\mathbf{G}\boldsymbol{p} + (\mathbf{A}\boldsymbol{p})^T\lambda^* \\
&= q(\boldsymbol{x}^*) + \frac{1}{2}\boldsymbol{u}^T\mathbf{Z}^T\mathbf{G}\mathbf{Z}\boldsymbol{u}
\end{aligned}
$$

$$q(\boldsymbol{x}) \quad = \quad q(\boldsymbol{x}^*) + \frac{1}{2}\boldsymbol{u}^T\mathbf{Z}^T\mathbf{G}\mathbf{Z}\boldsymbol{u}$$

Using the hypothesis that $\mathbf{Z}^T\mathbf{G}\mathbf{Z} \succeq 0$ we conclude that $q(\boldsymbol{x}) \geq q(\boldsymbol{x}^*)$ and $q(\boldsymbol{x}) = q(\boldsymbol{x}^*)$ iff $\boldsymbol{u} = 0$ otherwise $q(\boldsymbol{x}) > q(\boldsymbol{x}^*)$ and $\boldsymbol{x}^*$ is the unique global maximum.

Factoring the full KKT system

A general symmetric matrix $\mathbf{K}$ can be factorized (symmetric indefinite factorization, see Appendix Nocedal) as

$$\begin{aligned} \mathbf{P}^T\mathbf{K}\mathbf{P} &= \mathbf{L}\mathbf{B}\mathbf{L}^T \\ \mathbf{K} &= \mathbf{P}\mathbf{L}\mathbf{B}\mathbf{L}^T\mathbf{P}^T \end{aligned}$$

where $\mathbf{P}$ is a permutation matrix that introduce stability, $\mathbf{L}$ is unit lower triangular, and $\mathbf{B}$ is block-diagonal with either 1×1 or 2×2 blocks.
We want to solve

$$\left[\begin{array}{cc} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{array}\right]\left[\begin{array}{c} -\boldsymbol{p} \\ \lambda \end{array}\right] = \left[\begin{array}{c} \mathbf{g} \\ \mathbf{h} \end{array}\right] \tag{24}$$

Factoring the full KKT system

For solving $\mathbf{K}\boldsymbol{u} = \boldsymbol{v}$, we can group

$$\begin{aligned}\mathbf{P}^T\mathbf{K}\mathbf{P} &= \mathbf{L}\mathbf{B}\mathbf{L}^T \\ \mathbf{K}\boldsymbol{u} &= \mathbf{P}\mathbf{L}\left((\mathbf{B})\,\mathbf{L}^T\mathbf{P}^T\right)\boldsymbol{u} = \boldsymbol{v}\end{aligned}$$

- First let us define $\boldsymbol{z} = \left((\mathbf{B})\,\mathbf{L}^T\mathbf{P}^T\right)\boldsymbol{u}$ the we can solve $\mathbf{P}\mathbf{L}\boldsymbol{z} = \boldsymbol{v}$ or more specifically

$$\mathbf{L}\boldsymbol{z} = \mathbf{P}^T\boldsymbol{v}$$

- After obtaining $\boldsymbol{z}$ we have to solve the new system $\left((\mathbf{B})\,\mathbf{L}^T\mathbf{P}^T\right)\boldsymbol{u} = \boldsymbol{z}$

Factoring the full KKT system

- Now, let us define $\hat{\boldsymbol{z}} = \mathbf{L}^T\mathbf{P}^T\boldsymbol{u}$ then we can solve

$$\mathbf{B}\hat{\boldsymbol{z}} = \boldsymbol{z}$$

- After obtaining $\hat{\boldsymbol{z}}$ we have the system $\mathbf{L}^T\mathbf{P}^T\boldsymbol{u} = \hat{\boldsymbol{z}}$
- Defining $\bar{\boldsymbol{z}} = \mathbf{P}^T\boldsymbol{u}$ we can solve the system

$$\mathbf{L}^T\bar{\boldsymbol{z}} = \hat{\boldsymbol{z}}$$

- From $\mathbf{P}^T\boldsymbol{u} = \bar{\boldsymbol{z}}$ we can compute the solution by

$$\boldsymbol{u} = \mathbf{P}\bar{\boldsymbol{z}}$$

Factoring the full KKT system: summary

- For solving the KKT system we need to perform the following sequence of operations

$$
\begin{array}{rcl}
\mathbf{L}\boldsymbol{z} & = & \mathbf{P}^T\boldsymbol{v} \\
\mathbf{B}\hat{\boldsymbol{z}} & = & \boldsymbol{z} \\
\mathbf{L}^T\bar{\boldsymbol{z}} & = & \hat{\boldsymbol{z}} \\
\boldsymbol{u} & = & \mathbf{P}\bar{\boldsymbol{z}}
\end{array}
$$

where $\boldsymbol{v} = [\mathbf{g}; \mathbf{h}]$ and $[-\boldsymbol{p}; \lambda] = \boldsymbol{u}$ and $\mathbf{g} = \mathbf{G}\boldsymbol{x} + \boldsymbol{c}$, $\mathbf{h} = \mathbf{A}\boldsymbol{x} - \boldsymbol{b}$ and $\boldsymbol{x}^* = \boldsymbol{x} + \boldsymbol{p}$

Schur-complement method

Assuming that G is positive definite and from

$$
\begin{aligned}
-\mathbf{G}\boldsymbol{p} + \mathbf{A}^T\lambda &= \mathbf{g} \qquad (25)\\
-\mathbf{A}\boldsymbol{p} &= \mathbf{h} \qquad (26)
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T\lambda &= \mathbf{A}\mathbf{G}^{-1}\mathbf{g} - \mathbf{h} \qquad (27)\\
\mathbf{G}\boldsymbol{p} &= \mathbf{A}^T\lambda - \mathbf{g} \qquad (28)
\end{aligned}
$$

that requires the computation of the inverse of $\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T$ and $\mathbf{G}$. This is useful when $\mathbf{G}^{-1}$ is well conditioned and easy to invert or it is known explicitly through a quasi-Newton updating formula.

Schur-complement method

The name Schur-Complement method by applying block Gaussian elimination to

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \tag{29}$$

and using $\mathbf{G}$ as the pivot to obtain

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ 0 & -\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T \end{bmatrix} \tag{30}$$

following this procedure (block Gaussian elimination) we obtain the schur complement $-\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T$ of $\mathbf{G}$

Schur-complement method

Using the same strategy (block Gaussian elimination) for obtaining the schur complement, we can obtain the inverse of the matrix, ie

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} -\boldsymbol{p} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} \tag{31}$$

then

$$\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T\lambda = \mathbf{A}\mathbf{G}^{-1}\mathbf{g} - \mathbf{h} \tag{32}$$

$$\mathbf{G}\boldsymbol{p} = \mathbf{A}^T\lambda - \mathbf{g} \tag{33}$$

Schur-complement method

$$
\begin{aligned}
\lambda &= (\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}(\mathbf{A}\mathbf{G}^{-1}\mathbf{g} - \mathbf{h}) & (34)\\
\boldsymbol{p} &= \mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}(\mathbf{A}\mathbf{G}^{-1}\mathbf{g} - \mathbf{h}) - \mathbf{G}^{-1}\mathbf{g} & (35)
\end{aligned}
$$

Schur-complement method

$$
\begin{aligned}
-\boldsymbol{p} &= \mathbf{C}\mathbf{g} + \mathbf{E}\mathbf{h} && (36)\\
\lambda &= \mathbf{E}^T\mathbf{g} + \mathbf{F}\mathbf{h} && (37)
\end{aligned}
$$

with

$$
\begin{aligned}
\mathbf{C} &= \mathbf{G}^{-1} - \mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1} && (38)\\
\mathbf{E} &= \mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1} && (39)\\
\mathbf{F} &= -(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1} && (40)
\end{aligned}
$$

Schur-complement method

Finally

$$\begin{bmatrix} \mathbf{G} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C} & \mathbf{E} \\ \mathbf{E}^T & \mathbf{F} \end{bmatrix} \tag{41}$$

with

$$\mathbf{C} = \mathbf{G}^{-1} - \mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1} \tag{42}$$

$$\mathbf{E} = \mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1} \tag{43}$$

$$\mathbf{F} = -(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1} \tag{44}$$

Null-space method

- The null-space method does not require nonsingularity of $\mathbf{G}$ and therefore has wider applicability than the Schur-complement method. If that $\mathbf{A}$ has full row rank and that $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ is positive definite.
- Let us decompose $\boldsymbol{p}$ as

$$\boldsymbol{p} = \mathbf{Y}\boldsymbol{p}_Y + \mathbf{Z}\boldsymbol{p}_Z \tag{45}$$

where $[\mathbf{Y}, \mathbf{Z}]$ is a nonsingular matrix and $\mathbf{Z} \in \mathbb{R}^{n\times(n-m)}$ is a base of the null space of $\mathbf{A}$ and therefore $\mathbf{A}\mathbf{Z} = 0$

Null-space method

- From $\mathbf{A}\boldsymbol{p} = -\mathbf{h}$, $\mathbf{AZ} = 0$ and $\mathbf{AY}$ has inverse (see previous class)

$$\begin{aligned} \mathbf{A}\boldsymbol{p} &= \mathbf{AY}\boldsymbol{p}_Y + \mathbf{AZ}\boldsymbol{p}_Z = -\mathbf{h} && (46) \\ \mathbf{AY}\boldsymbol{p}_Y &= -\mathbf{h} && (47) \\ \boldsymbol{p}_Y &= -(\mathbf{AY})^{-1}\mathbf{h} && (48) \end{aligned}$$

- Then, from $-\mathbf{G}\boldsymbol{p} + \mathbf{A}^T\lambda = \mathbf{g}$ and $\mathbf{AZ} = 0$

$$\begin{aligned} -\mathbf{GY}\boldsymbol{p}_Y + \mathbf{GZ}\boldsymbol{p}_Z + \mathbf{A}^T\lambda &= \mathbf{g} \\ \mathbf{Z}^T\mathbf{GZ}\boldsymbol{p}_Z + \mathbf{Z}^T\mathbf{A}^T\lambda &= \mathbf{Z}^T(\mathbf{g} + \mathbf{GY}\boldsymbol{p}_Y) \\ \boldsymbol{p}_Z &= (\mathbf{Z}^T\mathbf{GZ})^{-1}\mathbf{Z}^T(\mathbf{g} + \mathbf{GY}\boldsymbol{p}_Y) \end{aligned}$$

Null-space method

- Finally

$$
\begin{aligned}
\boldsymbol{p}_Y &= -(\mathbf{A}\mathbf{Y})^{-1}\mathbf{h} &\quad (49)\\
\boldsymbol{p}_Z &= (\mathbf{Z}^T\mathbf{G}\mathbf{Z})^{-1}\mathbf{Z}^T(\mathbf{g}+\mathbf{G}\mathbf{Y}\boldsymbol{p}_Y) &\quad (50)
\end{aligned}
$$

- The Lagrange multiplier can be calculated as follows

$$
\begin{aligned}
-\mathbf{G}\boldsymbol{p}+\mathbf{A}^T\lambda &= \mathbf{g}\\
\mathbf{Y}^T\mathbf{A}^T\lambda &= \mathbf{Y}^T(\mathbf{g}+\mathbf{G}\boldsymbol{p})\\
\lambda &= (\mathbf{A}\mathbf{Y})^{-T}\mathbf{Y}^T(\mathbf{g}+\mathbf{G}\boldsymbol{p})
\end{aligned}
$$

Preconditioned CG for Reduced Systems

Let us express the solution of the quadratic program as

$$\boldsymbol{x} = \mathbf{Y}\boldsymbol{x}_Y + \mathbf{Z}\boldsymbol{x}_Z$$

Preconditioned CG for Reduced Systems

The general idea is to use elimination method followed of CG.

- Elimination

$$\begin{aligned}
\mathbf{A}\boldsymbol{x} &= \mathbf{A}\mathbf{Y}\boldsymbol{x}_Y + \mathbf{A}\mathbf{Z}\boldsymbol{x}_Z = \boldsymbol{b} \\
\mathbf{A}\mathbf{Y}\boldsymbol{x}_Y &= \boldsymbol{b} \\
\boldsymbol{x}_Y &= (\mathbf{A}\mathbf{Y})^{-1}\boldsymbol{b}
\end{aligned}$$

Preconditioned CG for Reduced Systems

Then

$$\boldsymbol{x} \quad = \quad \mathbf{Y}(\mathbf{AY})^{-1}\boldsymbol{b} + \mathbf{Z}\boldsymbol{x}_Z$$

and

$$\min_{\boldsymbol{x}_Z} \frac{1}{2}\boldsymbol{x}_Z\mathbf{Z}^T\mathbf{G}\mathbf{Z}\boldsymbol{x}_Z + \boldsymbol{x}_Z^T\boldsymbol{c}_Z$$

with

$$\boldsymbol{c}_Z \quad = \quad \mathbf{Z}^T\mathbf{G}\mathbf{Y}\boldsymbol{x}_Y + \mathbf{Z}^T\boldsymbol{c}$$

Preconditioned CG for Reduced Systems

We need to solve the following system

$$\mathbf{Z}^T\mathbf{G}\mathbf{Z}\boldsymbol{x}_Z = -\boldsymbol{c}_Z$$

Now, we can use GC to solve the previous system, due to $\mathbf{Z}^T\mathbf{G}\mathbf{Z}$ is positive definite.

Preconditioned CG for Reduced Systems

Require: $\boldsymbol{x}_0$

$\boldsymbol{g}_0 = \mathbf{Q}\boldsymbol{x}_0 - \boldsymbol{b}$,

Solve $\mathbf{M}\boldsymbol{z}_0 = g_0$ for $\boldsymbol{z}_0$

$\boldsymbol{d}_0 = -\boldsymbol{z}_0$, $k = 0$

while $\|\boldsymbol{g}_k\| \neq 0$ **do**

$\alpha_k = \frac{\boldsymbol{g}_k^T \boldsymbol{z}_k}{\boldsymbol{d}_k \mathbf{Q} \boldsymbol{d}_k}$

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{d}_k$

$\boldsymbol{g}_{k+1} = \boldsymbol{g}_k + \alpha_k \mathbf{Q}\boldsymbol{d}_k$

Solve $\mathbf{M}\boldsymbol{z}_{k+1} = \boldsymbol{g}_{k+1}$ for $\boldsymbol{z}_{k+1}$

$\beta_{k+1} = \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{z}_{k+1}}{\boldsymbol{g}_k^T \boldsymbol{z}_k}$

$\boldsymbol{d}_{k+1} = -\boldsymbol{z}_{k+1} + \beta_{k+1}\boldsymbol{d}_k$.

$k = k + 1$

end while

Preconditioned CG for Reduced Systems

- In the Previous algorithm we use the reduce Hessian, ie $\mathbf{Q} = \mathbf{Z}^T \mathbf{G} \mathbf{Z}$, and $\boldsymbol{b} = -\boldsymbol{c}_Z$
- This preconditioner $\mathbf{M}$ can be selected based on

$$\mathbf{M}^{-1/2}(\mathbf{Z}^T \mathbf{G} \mathbf{Z})\mathbf{M}^{-1/2} \quad = \quad \boldsymbol{I}$$

  with $\mathbf{M} = \mathbf{Z}^T \mathbf{G} \mathbf{Z}$
- As a preconditioner we can use $\mathbf{M} = \mathbf{Z}^T \mathbf{H} \mathbf{Z}$ where $\mathbf{H}$ is a positive definite matrix.
- For the selection of $\mathbf{H}$ we can use
  1. $\mathbf{H} = \mathbf{G}$
  2. $\mathbf{H} = \boldsymbol{I}$
  3. $\mathbf{H} = \text{diag}(|\mathbf{G}|)$

Projected CG Method