# Schur complement trick for positive semi-definite energies

Alec Jacobson Columbia University


## Abstract

The "Schur complement trick" appears sporadically in numerical optimization methods [Schur 1917; Cottle 1974]. The trick is especially useful for solving Lagrangian saddle point problems when minimizing quadratic energies subject to linear equality constraints [Gill et al. 1987]. Typically, to apply the trick, the energy's Hessian is assumed positive definite. I generalize this technique for positive *semi*-definite Hessians.


## 1 Positive definite energies

Let us consider a quadratic energy optimization problem subject to linear equality constraints:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \frac{1}{2}\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} - \mathbf{x}^{\mathsf{T}}\mathbf{f} + \text{constant}, \tag{1}$$

$$\text{subject to} \quad \mathbf{B}\mathbf{x} = \mathbf{g}, \tag{2}$$

where $\mathbf{x}, \mathbf{f} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times n}$ and $\mathbf{g} \in \mathbb{R}^m$.

Solving with the Lagrange multiplier method results in a system of linear equations:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^{\mathsf{T}} \\ \mathbf{B} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix},$$

where $\lambda \in \mathbb{R}^m$ is a vector of Lagrange multipliers.

To retain generality, let us replace the zero block in our system matrix with a variable $\mathbf{C}$.

$$\begin{bmatrix} \mathbf{A} & \mathbf{B}^{\mathsf{T}} \\ \mathbf{B} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix},$$

where $\lambda \in \mathbb{R}^m$ is a vector of Lagrange multipliers.

By assuming that $\mathbf{A}$ is positive definite, the Schur complement trick proceeds by multiplying the first set of equations by $\mathbf{B}\mathbf{A}^{-1}$:

$$\mathbf{B}\mathbf{A}^{-1}\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}}\lambda = \mathbf{B}\mathbf{A}^{-1}\mathbf{f}, \tag{3}$$

$$\mathbf{B}\mathbf{x} + \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}}\lambda = \mathbf{B}\mathbf{A}^{-1}\mathbf{f}. \tag{4}$$

Now, substitute the second set of equations $\mathbf{B}\mathbf{x} + \mathbf{C}\lambda = \mathbf{g}$ for $\mathbf{B}\mathbf{x}$ and solve the resulting equation for $\lambda$:

$$(\mathbf{g} - \mathbf{C}\lambda) + \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}}\lambda = \mathbf{B}\mathbf{A}^{-1}\mathbf{f}, \tag{5}$$

$$(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}} - \mathbf{C})\lambda = \mathbf{B}\mathbf{A}^{-1}\mathbf{f} - \mathbf{g}, \tag{6}$$

$$\lambda = (\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}} - \mathbf{C})^{-1}\left(\mathbf{B}\mathbf{A}^{-1}\mathbf{f} - \mathbf{g}\right). \tag{7}$$

Finally, find the primary solution by solving the first equation using the newly found values for $\lambda$:

$$\mathbf{x} = \mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^{\mathsf{T}}\lambda).$$

Assuming a factorization of $\mathbf{A}$ may be precomputed, this trick allows quickly solving optimization problems involving the same energy Hessian $\mathbf{A}$, but different linear coefficients $\mathbf{f}$ and different constraints $\mathbf{B}\mathbf{x} = \mathbf{g}$. So long as the number of constraints is significantly smaller than the number of variables ($m \ll n$), the cost of solving against the Schur complement ($\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{\mathsf{T}} - \mathbf{C}$) will be small compared to a full factorization (e.g. LDLT) of the system matrix $[\mathbf{A}\mathbf{B}^{\mathsf{T}};\mathbf{B}\mathbf{C}]$. This trick is beneficial in scenarios where the energy is fixed but a small number of constraints are changing.

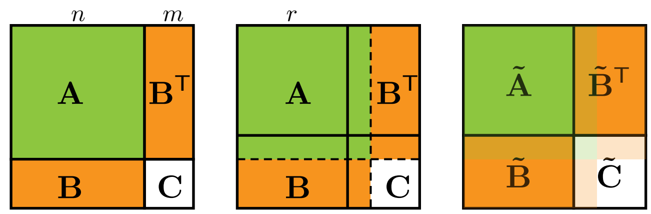


**Figure 1:** *If $\mathbf{A}$ is an $n \times n$ positive semi-definite matrix with rank $r$, then simply move $n-r$ rows and columns to the $\mathbf{B}$ and $\mathbf{C}$ blocks.*

## 2 Positive semi-definite energies

With loss of generality, assume $\mathbf{A}$ is symmetric, but merely positive semi-definite, with known rank $r < n$. We would like to apply the Schur complement trick from the previous section, but $\mathbf{A}$ is singular so we cannot factor it or solve against it.

However, we can simply shave off $n - r$ linearly independent rows and columns of $\mathbf{A}$ and push them into the $\mathbf{B}, \mathbf{B}^{\mathsf{T}}, \mathbf{C}$ blocks of the system system (see Figure 1). The remaining square portion of $\mathbf{A}$, $\tilde{A} \in \mathbb{R}^{r\times r}$, is a full rank and non-singular. Assuming the original system matrix $\mathbf{M} = [\mathbf{A}\mathbf{B}^{\mathsf{T}};\mathbf{B}\mathbf{C}] = [\tilde{\mathbf{A}}\tilde{\mathbf{B}}^{\mathsf{T}};\tilde{\mathbf{B}}\tilde{\mathbf{C}}]$ was non-singular, then new Schur complement ($\tilde{\mathbf{B}}\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{B}}^{\mathsf{T}} - \tilde{\mathbf{C}}$) will also be non-singular. This follows immediately from Schur's original observation that:

$$\det \mathbf{M} = \det \tilde{\mathbf{A}} \, \det(\tilde{\mathbf{B}}\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{B}}^{\mathsf{T}} - \tilde{\mathbf{C}}).$$

We can now simply apply the trick from the previous section.

This generalized trick is beneficial when the fixed energy has a non-trivial, but small null space.

Such situations arise in geometry processing when $\mathbf{A}$ is the Laplace or Laplace-Beltrami operator and the problem is minimizing Dirichlet energy subject to some *yet to be determined* boundary conditions or constraints. For a mesh with $n$ vertices, the discrete Laplace operator is rank $n - 1$, so only one row and column need to be moved.

One alternative to the presented approach would be to regularize $\mathbf{A}$ (e.g. $\mathbf{A} + \varepsilon\mathbf{I}$), but then one must choose between retaining the exact solution or ensuring numerical stability.

*I am grateful to Jernej Barbić, Keenan Crane, and Ladislav Kavan for illuminating discussions.*

## References


COTTLE, R. W. 1974. Manifestations of the schur complement. *Linear Algebra and its Applications 8*, 3.

GILL, P. E., MURRAY, W., SAUNDERS, M. A., AND WRIGHT, M. H. 1987. A schur-complement method for sparse quadratic programming. Tech. rep., Stanford University.

SCHUR, I. 1917. Über potenzreihen, die im innern des einheitskreises beschrankt sind. *Journal für die reine und angewandte Mathematik*.