Estimation of the HO for linear image reconstruction without explicit calculation of the reconstruction matrix

I. Definitions and goal

Consider the equation for the Hotelling Observer template

$$K_x w = s_x$$

where $K_x \in \Re^{n\times n}$ is the covariance matrix in the masked region of the reconstruction, $w \in \Re^n$ is the Hotelling template, and $s_x \in \Re^n$ is the mean masked, reconstructed signal. Denoting the reconstruction operator $R \in \Re^{p\times m}$ and the mask operator $M \in \Re^{n\times p}$, we have

$$K_x = MRK_bR^TM^T$$

where $K_b \in \Re^{m\times m}$ is the data domain covariance matrix (for an SKE/BKE task this matrix is diagonal and easy to calculate). The mean masked, reconstructed signal is given by

$$s_x = MRs_b$$

where $s_b \in \Re^m$ is the data domain signal (this is also easy to calculate).

We consider optimization-based reconstruction using a quadratic reconstruction optimization problem of the form

$$\min_x \|Ax - b\|^2 + \lambda\|Bx\|^2$$

where $A \in \Re^{p\times m}$ is the tomosynthesis forward model, $x \in \Re^p$ is the image estimate, $b \in \Re^m$ is the measured data, and $B \in \Re^{r\times p}$ is a regularization matrix (e.g. finite differencing). The solution to this optimization problem can be written

$$\begin{aligned} x^* &= (A^TA + \lambda B^TB)^{-1}A^Tb \\ &= Rb \end{aligned}$$

where we have defined

$$R = (A^TA + \lambda B^TB)^{-1}A^T$$

In practice, we implement A by calculating it's action on individual vectors (on-the-fly implementation); we do not store the matrix (it is very very large). Even if we could store it, evaluating the inverse of $(A^TA + \lambda B^TB)$ would be impractically time consuming. Our goal is therefore to somehow calculate the Hotelling template w without ever having to store or invert any large matrices.

II. Fussing around to pose the problem without inverses or storage of large matrices

We begin by plugging in the definition of K_x and s_x into the equation for the HO template

$$MRK_bR^TM^Tw = MRs_b$$

Define $Q = K_b^{\frac{1}{2}}R^TM^T$ (K_b is guaranteed to be positive semi-definite so it has a well-defined symmetric square root). The equation can the be written

$$Q^T(Qw - K_b^{-\frac{1}{2}}s_b) = 0$$

One may recognize this as the first order optimality condition (normal equation) of the least squares problem

$$\min_w \|Qw - K_b^{-\frac{1}{2}}s_b\|^2$$

or, plugging back in the definition of Q

$$\min_w \|K_b^{\frac{1}{2}}R^TM^Tw - K_b^{-\frac{1}{2}}s_b\|^2$$

One possibility would be to attempt to solve this problem with CG to obtain the HO template, which would involve performing a reconstruction at every iteration of CG. This may be feasible, but we would like to avoid having to perform a reconstruction in the inner loop of an iterative algorithm for obtaining the HO template.

To avoid having to calculate R we can do the following: first we plug in the definition of R to the least squares problem.

$$\min_{w} \|K_b^{\frac{1}{2}} A(A^T A + \lambda B^T B)^{-1} M^T w - K_b^{-\frac{1}{2}} s_b\|^2$$

We can equivalently repose the problem as

$$\begin{gathered}\min_{w,u} \|K_b^{\frac{1}{2}} Au - K_b^{-\frac{1}{2}} s_b\|^2 \\ \text{such that} \quad (A^T A + \lambda B^T B)u = M^T w\end{gathered}$$

The constraint states that $(A^T A + \lambda B^T B)u$ must lie in the row space of M, or equivalently, that $(A^T A + \lambda B^T B)u$ must be 0 outside the ROI. This can be rewritten equivalently as

$$\begin{gathered}\min_{u} \|K_b^{\frac{1}{2}} Au - K_b^{-\frac{1}{2}} s_b\|^2 \\ \text{such that} \quad (I - M^T M)(A^T A + \lambda B^T B)u = 0\end{gathered}$$

We have reached an optimization problem where everything involved can be calculated on the fly and there are no inverses or requirements of storage for large matrices. The unfortunate tradeoff is that we have traded an n-dimensional (ROI size) problem for a p-dimensional problem (reconstructed image size).

III. Equivalent symmetric indefinite system

Here we will show that a general linear equality constrained linear least squares (LSE) problem can be rewritten as a symmetric indefinite linear system. For now, ignore all variable definitions from the previous two sections as we will work more generally and then plug in our specific results at the end. We begin with the LSE problem

$$\min_{x} \|Qx - b\|^2 \text{ such that } Cx = g$$

The Lagrangian is

$$L(x, \mu) = \|Qx - b\|^2 + \mu^T (Cx - g)$$

Taking the gradient with respect to x and setting it equal to zero gives

$$Q^T Qx + C^T \mu = Q^T b$$

Combining this with the constraint equation we have

$$\begin{pmatrix} Q^T Q & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} Q^T b \\ g \end{pmatrix}$$

which is a symmetric indefinite linear system.

In our specific case this equation becomes

$$\begin{pmatrix} A^T K_b A & (A^T A + \lambda B^T B)(I - M^T M) \\ (I - M^T M)(A^T A + \lambda B^T B) & 0 \end{pmatrix} \begin{pmatrix} u \\ y \end{pmatrix} = \begin{pmatrix} A^T s_b \\ 0 \end{pmatrix}$$

My approach had been to solve this equation for u and y using MINRES, then to calculate $w = M(A^T A + \lambda B^T B)u$, which should be the Hotelling template. The problem was difficult to solve for small values of λ. This is likely a result of poor conditioning of the matrix. The approach did work for large values of λ.

I also tried solving the LSE problem directly with the method of deferred correction. This essentially involves solving the penalized least squares problem

$$\min_{x} \|Qx - b\|^2 + \alpha \|Cx - g\|^2$$

for successively larger values of α. I was able to replicate my results with MINRES but not much more.

IV. Another possible approach

The LSE problem can also be written as a larger symmetric indefinite system in which formation of $Q^T Q$ is avoided

$$\begin{pmatrix} 0 & Q^T & C^T \\ Q & -I & 0 \\ C & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ r \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ b \\ g \end{pmatrix}$$

I never tried solving this system, but avoiding formation of $Q^T Q$ explicitly may avoid certain numerical issues since the condition number of $Q^T Q$ is the square of the condition number of Q.