This document sumarises a compact approach to Gram-Schmidt pan-sharpening. Instead of applying the traditional Gram-Schmidt process, it describes the process using matrix inversions and projections as well as the concept of upsampling the multispectral data in a domain free from cross-band correlations. That data should be upsampled in such a domain makes intuitive sense because, in such a domain, all band operations can be applied independetly without introducing cross-band correlation distortions.

An important step of removing the mean value of all bands (i.e. all bands have mean value of zero) has been omitted with the assumption that in any practical implementation, these mean values can be subtracted prior to implementation and re-added post implementation.

1 Signal representation

Suppose we're given an N_b band multispectral dataset, with each band represented by an $m \times n$ image. Imagine that the set of samples in each image is flattened and the entire multispectral dataset is gathered into a "skinny" matrix with dimensions $N_s \times N_b$ where $N_s = mn$.

$$\mathbf{M} = \begin{bmatrix} \mathbf{b}_1 = \begin{bmatrix} b_1(1) \\ b_1(2) \\ b_1(3) \\ \vdots \\ b_1(N_s) \end{bmatrix} & \mathbf{b}_2 = \begin{bmatrix} b_2(1) \\ b_2(2) \\ b_2(3) \\ \vdots \\ b_2(N_s) \end{bmatrix} & \dots & \mathbf{b}_{N_b} = \begin{bmatrix} b_{N_b}(1) \\ b_{N_b}(2) \\ b_{N_b}(3) \\ \vdots \\ b_{N_b}(N_s) \end{bmatrix} \end{bmatrix} \tag{1}$$

One assumes that the data have been properly coregistered so that each index k (in the above $k \in \{1, 2, \dots, N_s\}$) corresponds to a physical point on the terrain and that each $b_m(k)$ is the band measurement for this point on the ground.

Now assume that we possess also a pan band that measures with a higher resolution than the multispectral bands and that it has also been sampled at a rate or N_f times the multispectral data sampling rate. Let this data set (which consists of only one image) be represented as a long column vector $\mathbf{p}_0$ with dimensions $N_f N_s \times 1$. Let some representation, or estimation, of this vector be given by the low-resolution (the same resolution as the multispectral imagery) vector $\mathbf{p}_m$ with dimensions $N_s \times 1$. Further, $\mathbf{p}_m$ is sampled at the same rate, and is co-registered to, $\mathbf{b}_k$. Later on, we will come back to a suitable choice for $\mathbf{p}_m$.

2 Signal covariance matrix

Let the covariance matrix for the matrix

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{p}_m & \mathbf{M} \end{bmatrix} \tag{2}$$

be given by

$$\mathbf{R}_0 = \mathbf{H}_0^T \mathbf{H}_0 \tag{3}$$

This matrix has dimensions $(N_b + 1) \times (N_b + 1)$.

3 Transformation of the multispectral data to remove pan covariance

We now introduce the projection of $\mathbf{M}$ that is perpendicular to $\mathbf{p}_m$

$$\mathbf{M}_\perp = \left(\mathbf{I}_{N_b} - \frac{\mathbf{p}_m \mathbf{p}_m^T}{\mathbf{p}_m^T \mathbf{p}_m} \right) \mathbf{M}. \tag{4}$$

This matrix has dimensions $N_s \times N_b$ and columns that are all orthogonal to the vector $\mathbf{p}_m$.

4 Transformation to whiten the transformed multispectral data

Define the $N_s \times (N_b + 1)$ dimensional matrix

$$\mathbf{H}_1 = \begin{bmatrix} \mathbf{p}_m & \mathbf{M}_\perp \end{bmatrix} \tag{5}$$

and consider

$$\mathbf{R}_1 = \mathbf{H}_1^T \mathbf{H}_1. \tag{6}$$

One finds that

$$\begin{aligned} \mathbf{R}_1^{-1} \mathbf{H}_1^T \mathbf{H}_1 &= \mathbf{I} \\ \mathbf{R}_1^{\frac{1}{2}} \mathbf{R}_1^{-1} \mathbf{H}_1^T \mathbf{H}_1 &= \mathbf{R}_1^{\frac{1}{2}} \\ \mathbf{R}_1^{\frac{1}{2}} \mathbf{R}_1^{-1} \mathbf{H}_1^T \mathbf{H}_1 \mathbf{R}_1^{-\frac{1}{2}} &= \mathbf{I} \quad , \\ \mathbf{R}_1^{-\frac{1}{2}} \mathbf{H}_1^T \mathbf{H}_1 \mathbf{R}_1^{-\frac{1}{2}} &= \mathbf{I} \\ (\mathbf{H}_1 \mathbf{R}_1^{-\frac{1}{2}})^T \mathbf{H}_1 \mathbf{R}_1^{-\frac{1}{2}} &= \mathbf{I} \end{aligned} \tag{7}$$

where $\mathbf{I}$ represents the $(N_b+1)\times(N_b+1)$ identity matrix. One observes that the $N_s \times (N_b+1)$ dimensional matrix $\mathbf{H}_1\mathbf{R}_1^{-\frac{1}{2}}$ has "whitened" columns.

Since $\mathbf{p}_m^T\mathbf{M}_\perp = \mathbf{0}$, one finds that

$$\mathbf{R}_1 = \mathbf{H}_1^T\mathbf{H}_1 = \begin{bmatrix} \mathbf{p}_m^T\mathbf{p}_m & \mathbf{0}^T \\ \mathbf{0} & \mathbf{M}_\perp^T\mathbf{M}_\perp \end{bmatrix}, \tag{8}$$

thus

$$\mathbf{R}_1^{-1} = (\mathbf{H}_1^T\mathbf{H}_1)^{-1} = \begin{bmatrix} \frac{1}{\mathbf{p}_m^T\mathbf{p}_m} & \mathbf{0}^T \\ \mathbf{0} & (\mathbf{M}_\perp^T\mathbf{M}_\perp)^{-1} \end{bmatrix}, \tag{9}$$

and

$$\mathbf{R}_1^{-\frac{1}{2}} = \begin{bmatrix} \frac{1}{|\mathbf{p}_m|} & \mathbf{0}^T \\ \mathbf{0} & (\mathbf{M}_\perp^T\mathbf{M}_\perp)^{-\frac{1}{2}} \end{bmatrix} \tag{10}$$

5 The Gram-Schmidt prescription

The previous two sections represent an implementation of the Gram-Schmidt process. Data have been transformed into a domain where bands are uncorrelated. The core concept of the Gram-Schmidt pan sharpening algorithm is to upsample the multispectral bands in this "whitened" domain and to replace the low-resolution pan band with the high resolution pan band. Data are then inverse transformed to recover an estimate of the high-resolution data in the original image domain. Thus the Gram-Schmidt pan-sharpening presecription is to:

1. compute $\mathbf{H}_1\mathbf{R}_1^{-\frac{1}{2}} = \left[\mathbf{p}_m/|\mathbf{p}_m| \quad \mathbf{M}_\perp(\mathbf{M}_\perp^T\mathbf{M}_\perp)^{-\frac{1}{2}}\right]$ and note that the first column is just the unit vector in the direction $\mathbf{p}_m$
2. upsample the remaining "whitened" multispectral columns
3. replace the first column (the normalized $\mathbf{p}_m$) with the fully measured and normalized $\mathbf{p}_0$ band
4. invert the whitening process to obtain the pan-sharpened image

6 Inverting the Gram-Schmidt transformation

Let the signal produced by the third step of the process (i.e. the upsampled data with the pan band replaced by the high res pan band) be given by the $N_f N_s \times (N_b + 1)$ dimensional matrix $\mathbf{W}_u$ Inversion of the Gram schmidt process can be thought of a re-introducting correlations in the covariance matrix so that one recovers a signal with the original covariance matrix $\mathbf{R}_0$. This is accomplished with the transformation $\mathbf{X}_u = \mathbf{W}_u \mathbf{R}_1^{\frac{1}{2}}$. This reverses the "whitening" operation of the first step.

Define the $(N_b + 1) \times 1$ column vector,

$$\mathbf{e}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{11}$$

and the $(N_b + 1) \times N_b$ matrix

$$\mathbf{I}_B = \begin{bmatrix} \mathbf{0}^T \\ \mathbf{I}_{N_b} \end{bmatrix}, \tag{12}$$

where $\mathbf{I}_{N_b}$ is the $N_b \times N_b$ identity matrix. To recover the $N_f N_s \times N_b$ pan-sharpened multispectral signal in the original domain (where the pan band is not orthogonal to the multispectral bands), we compute

$$\mathbf{Y}_u = \mathbf{X}_u \left[\mathbf{I} + \frac{\mathbf{e}_0 \mathbf{e}_0^T \mathbf{R}_1}{\mathbf{e}_0^T \mathbf{R}_0 \mathbf{e}_0} \right] \mathbf{I}_B \tag{13}$$

7 A suitable estimate for the low-resolution pan data

The low resolution pan data is represented in the literature as a weighted sum of the multispectral data,

$$\mathbf{p}_m = \mathbf{M}\mathbf{w} \tag{14}$$

where $\mathbf{w}$ is an $N_b \times 1$ column vector. We select for $\mathbf{w}$ the solution to

$$\min_{\mathbf{w}} \{|\mathbf{M}\mathbf{w} - \mathbf{p}_{ds}|\} \text{ such that } \mathbf{w} \geq 0 \tag{15}$$

where $\mathbf{p}_{ds}$ is a downsampled version of the full resolution pan image. If all the elements of $\mathbf{w} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{p}_{ds}$ are greater than zero, then this is the solution, and the representation of the pan image is given by

$$\mathbf{p}_m = \mathbf{M}(\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{p}_{ds} \tag{16}$$

In the case where some element of $\mathbf{w}$ is less than zero, one can iteratively solve 15 with the following approach. Let each element of w_i of $\mathbf{w}$ be given by a_i^2. Then If $\mathbf{y} = \mathbf{M}^T\mathbf{p}_{ds}$ and $\mathbf{R}_M = \mathbf{M}^T\mathbf{M}$, then

$$\mathbf{a}_{n+1} = \mathbf{a}_n - \mathbf{H}^{-1}(\mathbf{a}_n)\mathbf{d}(\mathbf{a}_n) \tag{17}$$

where

$$\begin{aligned} \mathbf{v}(\mathbf{a}) &= \mathbf{aa} \quad \text{... elementwise multiplication} \\ \mathbf{H}(\mathbf{a}) &= 4\text{diag}\{\mathbf{R}_M\mathbf{v}(\mathbf{a})\} + 8\text{diag}\{\mathbf{a}\}\mathbf{R}_M\text{diag}\{\mathbf{a}\}, \\ \mathbf{d}(\mathbf{a}) &= 4\text{diag}\{\mathbf{a}\}\left[\mathbf{R}_M\mathbf{v}(\mathbf{a}) - \mathbf{y}\right] \end{aligned} \tag{18}$$

and where, one could start the iteration with $\mathbf{a} = \sqrt{|\mathbf{w}_0|}$ (elementwise absolute value and square root) with $\mathbf{w}_0 = \mathbf{R}_M^{-1}\mathbf{y}$

Note that using $\mathbf{p}_m = \mathbf{M}\mathbf{w}$ means that the $(N_b + 1) \times (N_b + 1)$ matrix $\mathbf{R}_0$ has rank N_b. Special care thus has to be taken to ensure that inverses or the Gram-Schmidt process does not suffer from numerical instabilites. This can most easily be accomplished with diagonal-loading.