Application of Lagrange multipliers - Example

Consider the linear model

$$\begin{aligned} u_{k+1} &= u_k + 2v_k, \\ v_{k+1} &= v_k + 3u_k \end{aligned}$$

and suppose we make observations of $\tilde{u}_0, \tilde{u}_1$ of u at times t_0, t_1 respectively, each with error variance σ_o^2. We consider the data assimilation problem with no background term.

In this case we can write the system of equations as

$$\mathbf{x}_{k+1} = \mathbf{M}\mathbf{x}_k,$$

with

$$\mathbf{x} = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

Since we observe u the observation operator at each step is

$$\mathbf{H} = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

and the observation error covariance matrix $\mathbf{R}$ is the scalar σ_o^2.

Direct method

The cost function is given by

$$\begin{aligned} \mathcal{J}(\mathbf{x}_0) &= \frac{1}{2}\sum_{i=0}^{1}(\mathbf{H}_i\mathbf{x}_i - \mathbf{y}_i)^T\mathbf{R}^{-1}(\mathbf{H}_i\mathbf{x}_i - \mathbf{y}_i) \\ &= \frac{1}{2}\frac{(u_0 - \tilde{u}_0)^2}{\sigma_o^2} + \frac{1}{2}\frac{(u_1 - \tilde{u}_1)^2}{\sigma_o^2} \\ &= \frac{1}{2}\frac{(u_0 - \tilde{u}_0)^2}{\sigma_o^2} + \frac{1}{2}\frac{(u_0 + 2v_0 - \tilde{u}_1)^2}{\sigma_o^2} \end{aligned}$$

Then

$$\nabla\mathcal{J}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial \mathcal{J}}{\partial u} \\ \frac{\partial \mathcal{J}}{\partial v} \end{pmatrix}$$

So

$$\nabla\mathcal{J}(\mathbf{x}_0) = \begin{pmatrix} \sigma_o^{-2}(u_0 - \tilde{u}_0) + \sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1) \\ 2\sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1) \end{pmatrix}$$

Note that to apply this method we need to calculate n components of the gradient vector, where is n the length of the vector $\mathbf{x}$. This is very expensive for large n.

Adjoint method

From the adjoint equations we have

$$\begin{aligned}\boldsymbol{\lambda}_1 &= -\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_1 - \mathbf{y}_1)\\ \boldsymbol{\lambda}_0 &= \mathbf{M}^T\boldsymbol{\lambda}_1 - \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x}_0 - \mathbf{y}_0).\end{aligned}$$

Hence we have

$$\begin{aligned}\boldsymbol{\lambda}_1 &= -\begin{pmatrix}1\\0\end{pmatrix}\sigma_o^{-2}(u_1 - \tilde{u}_1)\\ &= -\begin{pmatrix}\sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1)\\ 0\end{pmatrix}.\end{aligned}$$

Then

$$\begin{aligned}\boldsymbol{\lambda}_0 &= -\begin{pmatrix}1 & 3\\ 2 & 1\end{pmatrix}\begin{pmatrix}\sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1)\\ 0\end{pmatrix} - \begin{pmatrix}1\\0\end{pmatrix}\sigma_o^{-2}(u_0 - \tilde{u}_0)\\ &= -\begin{pmatrix}\sigma_o^{-2}(u_0 - \tilde{u}_0) + \sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1)\\ 2\sigma_o^{-2}(u_0 + 2v_0 - \tilde{u}_1)\end{pmatrix}\end{aligned}$$

Hence we see that $\boldsymbol{\lambda}_0 = -\nabla\mathcal{J}(\mathbf{x}_0)$. However, in this case we have calculated all the components of the gradient vector with just one run of the adjoint model.