Construction of Gaussian Surrogate Process Using Numerical and Modeling Error Uncertainty

H. Aghakhani, A. K. Patra and E. Spiller

SIAM CSE 2015

March 8, 2015

Table of contents

Introduction

Governing Equations

$$U_t + F(U)_x + G(U)_y = S(U)$$

Where:

$$U = (h, hv_x, hv_y)^T$$
$$F = (hv_x, hv_x^2 + 0.5k_{ap}g_z h^2, hv_x hv_y)^T$$
$$G = (hv_y, hv_x v_y, hv_y^2 + 0.5k_{ap}g_z h^2)^T$$

Governing Equations Cont.

$$S = (0, S_x, S_y)^T$$

$$S_x = g_x h - \frac{V_x}{\sqrt{V_x^2 + V_y^2}} \left(g_z h + \frac{h V_x^2}{r_x} \right) \tan(\phi_{bed})$$

$$- h k_{ap} \mathrm{sgn} \left(\frac{\partial V_x}{\partial y} \right) \frac{\partial (g_z h)}{\partial y} \sin(\phi_{int})$$

$$S_y = g_y h - \frac{V_y}{\sqrt{V_x^2 + V_y^2}} \left(g_z h + \frac{h V_y^2}{r_y} \right) \tan(\phi_{bed})$$

$$- h k_{ap} \mathrm{sgn} \left(\frac{\partial V_y}{\partial x} \right) \frac{\partial (g_z h)}{\partial x} \sin(\phi_{int})$$

Adjoint Definition

Let:

- Let U and V be two vector spaces, and L be a linear operator that maps any $u \in U$ into $v \in V$.
- $< \cdot, \cdot >$ be a bilinear map that maps any two vectors like u, v two a real number, $U \times V \rightarrow \mathbb{R}$.

Then the adjoint operator, L^*, of L is defined:
$< Lu, v > = < u, L^* v >$.

Discrete Adjoint

Assume:

- $R(U, \alpha)$ as a system of governing equations,
- U is the solution vector,
- α is the vector of design parameters.

The object is to minimize $\quad \mathrm{J}(U, \alpha) \quad$ subject to $\quad R(U, \alpha) = 0$

Discrete Adjoint Cont.

With writing the variation of the functional and governing equation w.r.t design parameters, we will have:

$$\frac{dJ}{d\alpha} = \frac{\partial J}{\partial U}\frac{dU}{d\alpha} + \frac{\partial J}{\partial \alpha}$$

and:

Discrete Adjoint Cont.

With writing the variation of the functional and governing equation w.r.t design parameters, we will have:

$$\frac{dJ}{d\alpha} = \frac{\partial J}{\partial U}\frac{dU}{d\alpha} + \frac{\partial J}{\partial \alpha}$$

and:

$$\frac{\partial R}{\partial U}\frac{dU}{d\alpha} + \frac{\partial R}{\partial \alpha} = 0.$$

Discrete Adjoint Cont.

With writing the variation of the functional and governing equation w.r.t design parameters, we will have:

$$\frac{dJ}{d\alpha} = \frac{\partial J}{\partial U}\frac{dU}{d\alpha} + \frac{\partial J}{\partial \alpha}$$

and:

$$\frac{\partial R}{\partial U}\frac{dU}{d\alpha} + \frac{\partial R}{\partial \alpha} = 0.$$

Now,if we replace $\frac{dU}{d\alpha}$ from the second equation into the first equation, we have:

$$\frac{dJ}{d\alpha} = -\frac{\partial J}{\partial U}(\frac{\partial R}{\partial U})^{-1}\frac{\partial R}{\partial \alpha} + \frac{\partial J}{\partial \alpha}$$

Discrete Adjoint Cont.

With writing the variation of the functional and governing equation w.r.t design parameters, we will have:

$$\frac{dJ}{d\alpha} = \frac{\partial J}{\partial U}\frac{dU}{d\alpha} + \frac{\partial J}{\partial \alpha}$$

and:

$$\frac{\partial R}{\partial U}\frac{dU}{d\alpha} + \frac{\partial R}{\partial \alpha} = 0.$$

Now,if we replace $\frac{dU}{d\alpha}$ from the second equation into the first equation, we have:

$$\frac{dJ}{d\alpha} = -\frac{\partial J}{\partial U}\left(\frac{\partial R}{\partial U}\right)^{-1}\frac{\partial R}{\partial \alpha} + \frac{\partial J}{\partial \alpha}$$

Sensitivity Computation

Assume that n is the size of vector U, and m is the size of vector α:

$$dJ_{scalar} = \underbrace{-\left[\frac{\partial J}{\partial U}\right]_{1\times n}\left[\frac{\partial R}{\partial U}^{-1}\right]_{n\times n}\left[\frac{\partial R}{\partial \alpha}\right]_{n\times m} d\alpha_{m\times 1}}_{\text{scalar}} + \underbrace{\left[\frac{\partial J}{\partial \alpha}\right]_{1\times m} d\alpha_{m\times 1}}_{\text{scalar}}$$

Sensitivity Computation

Assume that n is the size of vector U, and m is the size of vector α:

$$dJ_{scalar} = \underbrace{-\left[\frac{\partial J}{\partial U}\right]_{1\times n}\left[\frac{\partial R}{\partial U}^{-1}\right]_{n\times n}\left[\frac{\partial R}{\partial \alpha}\right]_{n\times m} d\alpha_{m\times 1}}_{\text{scalar}} + \underbrace{\left[\frac{\partial J}{\partial \alpha}\right]_{1\times m} d\alpha_{m\times 1}}_{\text{scalar}}$$

Above sensitivity can be computed in two ways:

1. Forward mode: first computes $(\frac{\partial R}{\partial U})^{-1}\frac{\partial R}{\partial \alpha}$

Sensitivity Computation

Assume that n is the size of vector U, and m is the size of vector α:

$$dJ_{scalar} = \underbrace{-\left[\frac{\partial J}{\partial U}\right]_{1\times n}\left[\frac{\partial R}{\partial U}^{-1}\right]_{n\times n}\left[\frac{\partial R}{\partial \alpha}\right]_{n\times m} d\alpha_{m\times 1}}_{\text{scalar}} + \underbrace{\left[\frac{\partial J}{\partial \alpha}\right]_{1\times m} d\alpha_{m\times 1}}_{\text{scalar}}$$

Above sensitivity can be computed in two ways:

1. Forward mode: first computes $(\frac{\partial R}{\partial U})^{-1}\frac{\partial R}{\partial \alpha}$
2. Adjoint mode: first computes $\frac{\partial J}{\partial U}(\frac{\partial R}{\partial U})^{-1} \rightarrow (\frac{\partial R}{\partial U})^T v = \frac{\partial J}{\partial U}^T$, v is adjoint solution

Adjoint definition

Sensitivity Computation

Assume that n is the size of vector U, and m is the size of vector α:

$$dJ_{scalar} = \underbrace{-\left[\frac{\partial J}{\partial U}\right]_{1\times n}\left[\frac{\partial R}{\partial U}^{-1}\right]_{n\times n}\left[\frac{\partial R}{\partial \alpha}\right]_{n\times m} d\alpha_{m\times 1}}_{\text{scalar}} + \underbrace{\left[\frac{\partial J}{\partial \alpha}\right]_{1\times m} d\alpha_{m\times 1}}_{\text{scalar}}$$

Above sensitivity can be computed in two ways:

1. Forward mode: first computes $(\frac{\partial R}{\partial U})^{-1}\frac{\partial R}{\partial \alpha}$
2. Adjoint mode: first computes $\frac{\partial J}{\partial U}(\frac{\partial R}{\partial U})^{-1} \rightarrow (\frac{\partial R}{\partial U})^{T} v = \frac{\partial J}{\partial U}^{T}$, v is adjoint solution

Adjoint definition

Advantage of Adjoint:

If the number of design parameters are more than the objective functionals, then the computational cost of the adjoint is much lower than the forward method.

Sensitivity Computation

Assume that n is the size of vector U, and m is the size of vector α:

$$dJ_{scalar} = \underbrace{-\left[\frac{\partial J}{\partial U}\right]_{1\times n}\left[\frac{\partial R}{\partial U}^{-1}\right]_{n\times n}\left[\frac{\partial R}{\partial \alpha}\right]_{n\times m} d\alpha_{m\times 1}}_{\text{scalar}} + \underbrace{\left[\frac{\partial J}{\partial \alpha}\right]_{1\times m} d\alpha_{m\times 1}}_{\text{scalar}}$$

Above sensitivity can be computed in two ways:

1. Forward mode: first computes $(\frac{\partial R}{\partial U})^{-1}\frac{\partial R}{\partial \alpha}$
2. Adjoint mode: first computes
   $\frac{\partial J}{\partial U}(\frac{\partial R}{\partial U})^{-1} \rightarrow (\frac{\partial R}{\partial U})^{T} v = \frac{\partial J}{\partial U}^{T}$, v is adjoint solution

Adjoint definition

Advantage of Adjoint:

If the number of design parameters are more than the objective functionals, then the computational cost of the adjoint is much lower than the forward method.

Computing Adjoint

To solve adjoint equation we need Transpose of Jacobian Matrix:

$$\boxed{\left(\frac{\partial R}{\partial U}\right)^T v = \frac{\partial J}{\partial U}^T}$$

TITAN2D uses Godunov finite volume with Euler explicit time scheme, so the descritized form of equations are:

$$U_i^{n+1} = U_i^n - \frac{\triangle t}{\triangle x}\{F^n_{i+\frac{1}{2}} - F^n_{i-\frac{1}{2}}\} - \frac{\triangle t}{\triangle y}\{G^n_{i+\frac{1}{2}} - G^n_{i-\frac{1}{2}}\}$$

$$\left(\frac{\partial R}{\partial U}\right)^T_{m\times m} = K_{ij}$$

where m is the number of time steps, and each K_{ij} is a

Computing Adjoint Cont.

For Euler explicit:

$$K_{i,i} = I \quad \text{and} \quad K_{i,i+1} = (\frac{\partial R_p^{i+1}}{\partial U_q^i})^T,$$

- p and q are degrees of freedom
- rest of the components are zero
- depend on the stencil is used. K matrices are also block bounded

$$(\frac{\partial R}{\partial U})^T_{m\times m} = \begin{pmatrix} I & K_{1,2} & & & \\ & I & K_{2,3} & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & I & K_{m-1,m} \\ & & & & I \end{pmatrix}$$

Computing Adjoint Cont.

Important conclusion:

To compute adjoint for an explicit system, there is no need to solve a system of equation, and adjoint solution can be found marching backward in time.

$$v_1 + K_{1,2} v_2 = \left(\frac{\partial J}{\partial U}\right)_1^T$$

$$\vdots$$

$$v_{m-1} + K_{m-1,m} v_m = \left(\frac{\partial J}{\partial U}\right)_{m-1}^T$$

$$v_m = \left(\frac{\partial J}{\partial U}\right)_m^T$$

Error Estimation

The goal is to minimize the numerical error due to mesh size on the objective functional $J(Q)$, given the solution on the coarse mesh $J(Q_H)$.
With Taylor expansion we can write[1]:

$$J(Q_h) \approx J(Q_h^H) - \underbrace{(\psi_h^H)^T R(Q_h^H)}_{\text{Adjoint correction term}} - \underbrace{(\psi_h - \psi_h^H)^T R(Q_h^H)}_{\text{Remaining term}},$$

where $J(Q_h)$ is the functional value on a finer mesh, and all $\square_h^H$ is the projection of $\square$ from the coarse mesh to the fine mesh.

[1] Marian Nemec, MJ Aftosmis, and Mathias Wintzer. "Adjoint-based adaptive mesh refinement for complex geometries". In: *AIAA Paper* (2008), pp. 1–23. URL: http://people.nas.nasa.gov/~nemec/MYWeb/aiaa%5C_2008%5C_0725%5C_small.pdf.

Error Estimation Cont.

So to compute the error $\varepsilon = |J(Q) - J(Q_H)|$:

- we have to first find ψ_h.
- since computing ψ_h is not reasonable to just compute the error we approximate it with higher order construction.

Consequently:

$$J(Q_h) \approx J(Q_L) - (\psi_L)^T R(Q_L) - (\psi_L - \psi_C)^T R(Q_L),$$

where $\Box_L$, $\Box_C$ represent linear and constant reconstruction respectively.

Sensitivity Computation

Sensitivity Computation

Sensitivity Computation

Connection to Adjoint definition

$$
\begin{aligned}
u = \frac{dU}{d\alpha}, \qquad A &= \frac{\partial R}{\partial U} \\
g^T = \frac{\partial J}{\partial U}, \qquad f &= -\frac{\partial R}{\partial \alpha}
\end{aligned} \tag{1}
$$

Forward method:

$$
\begin{gathered}
\frac{dJ}{d\alpha} = g^T u + \frac{\partial J}{\partial \alpha} \\
\text{Subject to} \quad Au = f
\end{gathered} \tag{2}
$$

Adjoint Method:

$$
\begin{gathered}
\frac{dJ}{d\alpha} = v^T f + \frac{\partial J}{\partial \alpha} \\
\text{Subject to} \quad A^T v = g
\end{gathered} \tag{3}
$$

Back to Connection adjoint definition.