Keith —
from St. Pet. / Jarny

Solution of a Heat Conduction Parameter Estimation Problem
Using The Adjoint Method(-25 Aug. 1994)

As interpreted and modified by J.V. Beck, MSU (oct. 1994)
and "revisited" by Y.C.Jarny, ISITEM (Dec.1994)

The parameter estimation problem in transient heat conduction for estimating constant thermal conductivity is considered using the adjoint method. The problem can be mathematically stated as

$$
\begin{aligned}
&k\frac{\partial^2 T}{\partial x^2} = C\frac{\partial T}{\partial t} \quad 0 < x < L \\
&-k\frac{\partial T}{\partial x}_{|x=0} = q(t) \qquad\qquad (1\text{-}4) \\
&\frac{\partial T}{\partial x}_{|x=L} = 0 \\
&T(x,0) = T_0(x), 0 < x < L
\end{aligned}
$$

where the thermal conductivity k and the volumetric heat capacity C are unknowns. There is some excess information of transient temperature measurements at locations $x_1, x_2, \ldots, x_N$; these measured temperatures over the time interval (o, t_f) are denoted $Y_1(t), Y_2(t), \ldots, Y_N(t)$. Additive measurement errors are assumed,

$$Y_j(t) = T(x_j, t) + \varepsilon_j(t) \qquad (5)$$

where $\varepsilon_j(t)$ is a measurement error withe zero mean, constant variance and is uncorrelated.

For convenience, let β be the vector

$$\beta = \begin{bmatrix} k \\ C \end{bmatrix} \qquad (6)$$

let β_n be the nth iterate and β_0 the initial value of β. Also we define a correction $\delta\beta_n$

$$\beta_{n+1} = \beta_n + \delta\beta_n \qquad (7)$$

For the above problem of additive errors, a sum of squares function $J(\beta)$ can be minimized to estimate β

$$J(\beta) = \sum_{j=1}^{N} \int_0^{t_f} \left[T(x_j, t; \beta) - Y_j(t)\right]^2 dt \quad (8)$$

Now in order to have the minimizing sequence

$$J(\beta_{n-1}) < J(\beta_n), \forall n$$

the correction $\delta\beta_n$ is to be found using gradient methods, for example:

$$\delta\beta_n = -\rho_n \nabla J(\beta_n) \qquad (9)$$

where ρ_n is a positive scalar and $\nabla J(\beta_n)$ is the gradient of $J(\beta)$ and has components ∇J_k and ∇J_C .The value of ρ_n can be found using steepest descent as it is discussed at the end. Conjugate gradient or other methods can be used.

In order to evaluate $\nabla J(\beta_n)$, we first introduce δT the variation of temperature and δJ the variation of $J(\beta)$ due to the correction $\delta\beta_n$. These quantities also called differential, are evaluated at β_n and they are given respectively by

$$\delta T(x,t;\beta_n) = \lim_{\mu\to 0}\frac{T(x,t;\beta_n+\mu\delta\beta_n)-T(x,t;\beta_n)}{\mu} \quad \textbf{(10)}$$

$$\delta J(\beta_n) = \lim_{\mu\to 0}\frac{J(\beta_n+\mu\delta\beta_n)-J(\beta_n)}{\mu} \quad \textbf{(11)}$$

Then using the adjoint equation, this differential δJ will be put in the form of a scalar product of two vectors:

$$\delta J(\beta_n) = \langle \nabla J(\beta_n), \delta\beta_n \rangle \quad \textbf{(12a)}$$

For the case of 2 constant parameters (k and C), this scalar product takes the standard form

$$\delta J(\beta_n) = \sum_{i=1}^{2}\frac{\partial J}{\partial \beta_i}\delta\beta_i^{(n)} = \frac{\partial J}{\partial k}\delta k^{(n)} + \frac{\partial J}{\partial C}\delta C^{(n)} \quad (12b)$$

Let us denote $T_\mu = T(x,t;\beta_n+\mu\delta\beta_n)$ the solution of eq.(1)-(4) evaluated at $\beta_n+\mu\delta\beta_n$..

$$\begin{aligned}
&(k+\mu\delta k)\frac{\partial^2 T_\mu}{\partial x^2} = (C+\mu\delta C)\frac{\partial T_\mu}{\partial t} \quad 0<x<L\\
&-(k+\mu\delta k)\left.\frac{\partial T_\mu}{\partial x}\right|_{x=0} = q(t)\\
&\left.\frac{\partial T_\mu}{\partial x}\right|_{x=L} = 0\\
&T_\mu(x,0) = T_0(x)
\end{aligned} \quad \textbf{(13)}$$

Substracting eq.(1-4) from eq.(13) and dividing by μ gives

$$\begin{aligned}
&k\frac{\partial^2[(T_\mu-T)/\mu]}{\partial x^2} = C\frac{\partial[(T_\mu-T)/\mu]}{\partial t} + \left[-\delta k\frac{\partial^2 T_\mu}{\partial x^2} + \delta C\frac{\partial T_\mu}{\partial t}\right] 0<x<L\\
&-k\left.\frac{\partial[(T_\mu-T)/\mu]}{\partial x}\right|_{x=0} = \delta k\left.\frac{\partial T_\mu}{\partial x}\right|_{x=0}\\
&\left.\frac{\partial[(T_\mu-T)/\mu]}{\partial x}\right|_{x=L} = 0\\
&\frac{T_\mu-T}{\mu}(x,0) = 0
\end{aligned} \quad \textbf{(14)}$$

Using the definition given by eq. (10), eq.(14) become:

$$k\frac{\partial^2 \delta T}{\partial x^2} = C\frac{\partial \delta T}{\partial t} + \left[-\delta k\frac{\partial^2 T}{\partial x^2} + \delta C\frac{\partial T}{\partial t}\right], 0 < x < L$$

$$-k\frac{\partial \delta T}{\partial x}_{|x=0} = \delta k\frac{\partial T}{\partial x}_{|x=0} \qquad \textbf{(15)}$$

$$\frac{\partial \delta T}{\partial x}_{|x=L} = 0$$

$$\delta T(x,0) = 0\ , 0 < x < L$$

Equations (15) are called sensitivity equations; notice that δT is the sensitivity for both k and C being varied simultaneously. But it is possible to consider the variation δT for just one parameter by taking the variation of the other parameter equal to zero.

Now we are able to calculate δJ according to the definition (12):

$$J(\beta_n + \mu\delta\beta_n) - J(\beta_n) = \sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta + \mu\delta\beta_n) - Y_j(t)\right]^2 dt - \sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta_n) - Y_j(t)\right]^2 dt$$

$$\frac{J(\beta_n + \mu\delta\beta_n) - J(\beta_n)}{\mu} = \sum_{j=1}^{N}\int_0^{t_f}\left[(T_\mu - T)/\mu\right]_{x=x_j}\left[(T_\mu + T)_{x=x_j} - 2Y_j(t)\right]dt$$

$$\delta J(\beta_n) = -2\sum_{j=1}^{N}\int_0^{t_f}\delta T(x_j,t;\beta_n)e_j^{(n)}(t)dt \qquad \textbf{(16)}$$

where $e_j^{(n)}(t)$ the residual is given by

$$e_j^{(n)}(t) = Y_j(t) - T(x_j,t;\beta_n) \qquad (17)$$

Notice that eq.(16) can also be written as

$$\delta J(\beta_n) = -2\sum_{j=1}^{N}\int_0^{t_f}\int_0^{L}\delta T(x,t;\beta_n)e_j^{(n)}(t).\delta(x - x_j)dxdt \qquad \textbf{(18)}$$

The adjoint equation is now derived using the method of Lagrange multipliers. The main idea of this method is based on the fact that the solution T of eq.(1)-(4) is dependent on the parameter β, these equations can be considered as constraints for the optimization of the criterion $J(\beta)$. So let us introduce the so-called Lagrangian multiplier ψ and the Lagrangian associated to the criterion $J(\beta)$ and the constraints (1)-(4):

$$\Lambda(\psi,T,\beta) = J(T) + \int_0^{t_f}\int_0^{L}\psi(x,t)\left[k\frac{\partial^2 T}{\partial x^2} - C\frac{\partial T}{\partial t}\right]dxdt \qquad (19)$$

Integrations by parts , twice in space, once in time in the second term, give an equivalent form of the Lagrangian:

$$\Lambda(\psi,T,\beta)=J(T)+\int_0^{t_f}\int_0^L T(x,t)\left[k\frac{\partial^2\psi}{\partial x^2}+C\frac{\partial\psi}{\partial t}\right]dxdt$$
$$+\int_0^{t_f}\left[k\psi\frac{\partial T}{\partial x}-k\frac{\partial\psi}{\partial x}T\right]_{x=L}dt-\int_0^{t_f}\left[k\psi\frac{\partial T}{\partial x}-k\frac{\partial\psi}{\partial x}T\right]_{x=0}dt \quad \textbf{(20)}$$
$$+\int_0^L\left[-CT\psi\right]_{t=t_f}dt+\int_0^L\left[CT\psi\right]_{t=0}dt$$

This result is valid for any function T, especially in the particular case where T satisfies the constraints given by eq.(1-4), let us denote $T(\beta)$ this particular function, then we have the Lagrangian equal to the criterion

$$\Lambda(\psi,T(\beta),\beta)=J(\beta) \qquad \textbf{(21)}$$

And assuming that the multiplier ψ is fixed, the differential of Λ resulting of the variation of T and β is equal to the differential of J

$$\delta\Lambda(\psi,T(\beta),\beta)=\delta J(\beta) \qquad \textbf{(22)}$$

The variation $\delta\Lambda$ due to arbitrary variations δT and $\delta\beta$ can be evaluated following the preceding rules, it comes from eq.(19) and (20)

$$\delta\Lambda(\psi,T,\beta)=\delta J+\int_0^{t_f}\int_0^L\delta T\left[k\frac{\partial^2\psi}{\partial x^2}+C\frac{\partial\psi}{\partial t}\right]dxdt$$
$$+\int_0^{t_f}\left[k\psi\frac{\partial\delta T}{\partial x}-k\frac{\partial\psi}{\partial x}\delta T\right]_{x=L}dt-\int_0^{t_f}\left[k\psi\frac{\partial\delta T}{\partial x}-k\frac{\partial\psi}{\partial x}\delta T\right]_{x=0}dt \quad \textbf{(23)}$$
$$+\int_0^L\left[-C\delta T\psi\right]_{t=t_f}dt+\int_0^L\left[C\delta T\psi\right]_{t=0}dt$$
$$+\int_0^{t_f}\int_0^L\psi\left[\delta k\frac{\partial^2 T}{\partial x^2}-\delta C\frac{\partial T}{\partial t}\right]dxdt$$

By putting the resulting differential δJ of eq.(18) in eq. (23), the first two terms of that equation become

$$\delta\Lambda(\psi,T,\beta)=\int_0^{t_f}\int_0^L\delta T\left[k\frac{\partial^2\psi}{\partial x^2}+C\frac{\partial\psi}{\partial t}-2\sum_{i=1}^{N}e_i\delta(x-x_i)\right]dxdt$$
$$+\int_0^{t_f}\left[k\psi\frac{\partial\delta T}{\partial x}-k\frac{\partial\psi}{\partial x}\delta T\right]_{x=L}dt-\int_0^{t_f}\left[k\psi\frac{\partial\delta T}{\partial x}-k\frac{\partial\psi}{\partial x}\delta T\right]_{x=0}dt \quad \textbf{(24)}$$
$$+\int_0^L\left[-C\delta T\psi\right]_{t=t_f}dt+\int_0^L\left[C\delta T\psi\right]_{t=0}dt$$
$$+\int_0^{t_f}\int_0^L\psi\left[\delta k\frac{\partial^2 T}{\partial x^2}-\delta C\frac{\partial T}{\partial t}\right]dxdt$$

In the particular case of our interest $T=T(\beta)$, by using the properties of $\delta T(x,t)$ given in the sensitivity eq.(15), we get

$$\delta\Lambda(\psi,T(\beta),\beta) = \int_0^{t_f}\int_0^L \delta T\left[k\frac{\partial^2\psi}{\partial x^2} + C\frac{\partial\psi}{\partial t} - 2\sum_{i=1}^{N} e_i\delta(x-x_i)\right]dxdt$$
$$+\int_0^{t_f}\left[-k\frac{\partial\psi}{\partial x}\delta T\right]_{x=L}dt - \int_0^{t_f}\left[-\delta k\frac{\partial T}{\partial x}\psi - k\frac{\partial\psi}{\partial x}\delta T\right]_{x=0}dt + \int_0^L\left[-C\delta T\psi\right]_{t=t_f}dt \qquad (25)$$
$$+\int_0^{t_f}\int_0^L \psi\left[\delta k\frac{\partial^2 T}{\partial x^2} - \delta C\frac{\partial T}{\partial t}\right]dxdt$$

Now for convenience, let us fixe the multiplier ψ as follows:

$$\begin{aligned} &k\frac{\partial^2\psi}{\partial x^2} + C\frac{\partial\psi}{\partial t} - 2\sum_{i=1}^{N} e_i\delta(x-x_i), 0<x<L \\ &k\frac{\partial\psi}{\partial x}\Big|_{x=0} = 0 \\ &k\frac{\partial\psi}{\partial x}\Big|_{x=L} = 0 \\ &\psi(x,t_f) = 0, 0<x<L \end{aligned} \qquad (26)$$

then the differential $\delta\Lambda(\psi,T(\beta),\beta)$ reduces to the final following form

$$\delta\Lambda(\psi,T(\beta),\beta) = \int_0^{t_f}\int_0^L \psi(x,t)\left[\delta k\frac{\partial^2 T}{\partial x^2} - \delta C\frac{\partial T}{\partial t}\right]dxdt - \int_0^{t_f}\left[\delta k\frac{\partial T}{\partial x}\psi\right]_{x=0}dt \qquad (27)$$

using the fact that the parameters are constant, and using the property given by eq: (22), the eq.(27) can be written in the desired form of the eq. (12)

$$\delta J(\beta) = -\delta C\int_0^{t_f}\int_0^L \frac{\partial T}{\partial t}\psi(x,t)dxdt - \delta k\int_0^{t_f}\int_0^L \frac{\partial\psi}{\partial x}\frac{\partial T}{\partial x}dxdt \qquad (28)$$

thus in conclusion of this part, the components of the gradient are related to adjoint variable by

$$\begin{aligned} \frac{\partial J}{\partial C} &= -\int_0^{t_f}\int_0^L \frac{\partial T}{\partial t}\psi(x,t)dxdt \\ \frac{\partial J}{\partial k} &= -\int_0^{t_f}\int_0^L \frac{\partial\psi}{\partial x}\frac{\partial T}{\partial x}dxdt \end{aligned} \qquad (29)$$

Since $T(x,t)$ and $\psi(x,t)$ are known, these components of the gradient can be evaluated. It can be observed that the first component is equal to zero for steady state, and that the second component is equal to zero without gradient of temperature in the body.

The direction for improving estimates of k and C are known but the size of the step needs to be found. In the method of steepest descent, we said eq.(9) that

$$\delta\beta = -\rho\nabla J(\beta) \qquad (30)$$

Then it is possible to choose the scalar ρ by considering the function $\phi(\rho)$ (which changes from iteration to iteration) and searching for the positive value $\hat{\rho}$ which mimimizes it

$$\phi(\rho) = J(\beta_n - \rho\nabla J(\beta_n)) \qquad \textbf{(31)}$$

The optimal $\hat{\rho}$ value satisfies the characteristic equation

$$\left.\frac{d\phi(\rho)}{d\rho}\right|_{\rho=\hat{\rho}} = \left.\frac{dJ(\beta_n - \rho\nabla J(\beta_n))}{d\rho}\right|_{\rho=\hat{\rho}} = 0 \qquad \textbf{(32)}$$

Note that the calculated temperature $T(x,t;\beta_n - \rho\nabla J(\beta_n))$ is not linear with the parameter $\beta_n - \rho\nabla J(\beta_n)$, but linearization can be done around the solution $T(x,t;\beta_n)$. Remember that the sensitivity eq.(15) gives the resulting variation $\delta T(x,t;\beta_n)$ in temperature due to a variation $\delta\beta$ of the parameter. So let us denote $\theta(x,t)$ the variation in temperature calculated with a variation of the parameter equal to $\nabla J(\beta_n)$, it is calculated from the adapted sensitivity equations:

$$\begin{aligned}
&k\frac{\partial^2\theta}{\partial x^2} = C\frac{\partial\theta}{\partial t} + \left[-\nabla J_k\frac{\partial^2 T}{\partial x^2} + \nabla J_C\frac{\partial T}{\partial t}\right], 0 < x < L \\
&-k\left.\frac{\partial\theta}{\partial x}\right|_{x=0} = \nabla J_k\left.\frac{\partial T}{\partial x}\right|_{x=0} \\
&\left.\frac{\partial\theta}{\partial x}\right|_{x=L} = 0 \\
&\theta(x,0) = 0 \;, 0 < x < L
\end{aligned} \qquad \textbf{(33)}$$

then we have the approximated new temperature

$$T(x,t;\beta_n - \rho\nabla J(\beta_n)) \approx T(x,t;\beta_n) - \rho\theta(x,t) \qquad \textbf{(34)}$$

By putting this approximation in the function $\phi(\rho)$ given by eq. (31), we get from the definition of eq.(8):

$$\phi(\rho) \approx \sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta) - \rho\theta(x_j,t) - Y_j(t)\right]^2 dt \qquad \textbf{(35)}$$

The development of eq. (35) gives a quadratic form versus the variable ρ:

$$\begin{aligned}
\phi(\rho) \approx\ & \rho^2\sum_{j=1}^{N}\int_0^{t_f}\left[\theta(x_j,t)\right]^2 dt - 2\rho\sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta_n)) - Y_j(t)\right]\theta(x_j,t)dt \\
& + \sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta_n)) - Y_j(t)\right]^2 dt
\end{aligned} \qquad \textbf{(36)}$$

Take the derivative of $\phi(\rho)$ with respect to ρ and set the expression equal to zero,

$$\left.\frac{d\phi(\rho)}{d\rho}\right|_{\rho=\hat{\rho}} = 2\hat{\rho}\sum_{j=1}^{N}\int_0^{t_f}\left[\theta(x_j,t)\right]^2 dt - 2\sum_{j=1}^{N}\int_0^{t_f}\left[T(x_j,t;\beta_n)) - Y_j(t)\right]\theta(x_j,t)dt \qquad \textbf{(37)}$$

Solving for $\hat{\rho}$ gives

$$\hat{\rho} = \frac{\sum_{j=1}^{N} \int_0^{t_f} \left[T(x_j, t; \beta_n)) - Y_j(t) \right] \theta(x_j, t) dt}{\sum_{j=1}^{N} \int_0^{t_f} \left[\theta(x_j, t) \right]^2 dt} \quad \textbf{(38)}$$

Now the solution of the Heat Conduction Parameter Estimation using the combination of the Adjoint method and the steepest descent method, can be put in the form of the following algorithm:

a.- Select the initial values $k^{(0)}$ and $C^{(0)}; n \leftarrow 0$

b.- $n \leftarrow n+1$

c.- Solve the direct problem given by eqs. (1-4) to get $T(x,t;\beta_n)$, the residuals $e_j^{(n)}(t), j = 1,..N$, eq. (17) and the criterion $J(\beta_n)$, eq.(8)

d.- Solve the adjoint problem given by eq.(26) to get $\psi(x,t)$ and the components of the gradient $\nabla J_k, \nabla J_C$, eq.(29)

e.- Solve the sensitivity problem given by eq.(33) to get $\theta(x,t)$ and the optimal step $\hat{\rho}$ eq.(38)

f.- Get an improved value of β eq.(8-9)

$$\beta_{n+1} = \beta_n - \hat{\rho} \nabla J(\beta_n)$$

g.- If the changes in the criterion (*) $J(\beta)$ or the changes in the parameter β are not small, goto step b, else

h.- end of the algorithm.
(*) Another approach is to use the residual principle.