MULTIWAVELENGTH PYROMETRY

The basic idea of multiwavelength pyrometry consists in the supposition that in a certain wavelength interval $[\lambda_{j_a} \dots \lambda_{j_b}]$, such that for the natural numbers j_b and j_a: $j_b - j_a = K - 1$ (i.e., K measured points of the spectrum get within this wavelength interval); a function with a limited $(L - 1)$ set of variable coefficients (a_l), where $l = 0 \dots L - 2$ can be used to approximate the emissivity. Then, provided that $K > L$, one can formulate the optimization problem of minimizing the residual between the calculated thermal emission spectrum $I_{sam,,ij}$ in the wavelength interval $[\lambda_{j_a} \dots \lambda_{j_b}]$, and its approximation in the space of L optimization variables $(a_0 \dots a_{L-2}, a_{L-1})$, where additional $(L - 1)$'th optimization variable is the true temperature $a_{L-1} = T_{true}$. For example, a polynomial approximation of the emissivity viz $\varepsilon_{sam,j} = \sum_{l=0}^{L-2} a_l \lambda^l{}_j$ can be used. The residual (r_j) between the experiment and calculation at wavelength $\lambda_j \in [\lambda_{j_a} \dots \lambda_{j_b}]$ can be introduced as follows:

$$r_j = y_j - M_j\left(\lambda_j, a_0 \dots a_{L-2}, a_{L-1}\right) \qquad \text{(A0)}$$

where $M_j = \left(I_{bb}\left(\lambda_j, a_{L-1}\right) - I_{sur,j}\right) \cdot \sum_{l=0}^{L-2} a_l \lambda^l{}_j$, and y_j – measured spectral intensity.

It is also convenient to introduce a new index k that runs from 1 to K, so that $\lambda_{k=1:K} = [\lambda_{j_a} \dots \lambda_{j_b}]$, index l of the optimization variables runs from 0 to $L - 1$. We can rewrite the quadratic residual minimization problem in a vector form:

$$\text{argmin}\left(f(\mathbf{x})\right), \text{ where } f(\mathbf{x}) = \|\mathbf{r}\|_2 = \frac{1}{2}\mathbf{r}^T\mathbf{r} = \frac{1}{2}\sum_{k=1}^{K} r_k{}^2, \qquad \text{(A1)}$$

where $\mathbf{x}$ is a vector-column of the optimization variables $\mathbf{x}^T = (a_0 \dots a_{L-2}, a_{L-1}) = (\mathbf{a}^T, T_{true})$, the superscript T denotes transposition, and $\mathbf{r} = \mathbf{y} - \mathbf{M}(\mathbf{x})$ is the expression (8) written in a vector form, $\mathbf{a}$ is a column-vector of polynomial coefficients approximating emissivity. Further in this section we will assume that vectors (designated as bolded symbols) mean vector columns, K is the number of

spectral points, and L is the number of the optimization variables. The necessary conditions that a vector $\mathbf{x}^*$ is the local minimizer of the problem (A1) are $\nabla f(\mathbf{x}^*) = \mathbf{0}, \nabla^2 f(\mathbf{x}^*) \geq 0$ (i.e., the Hessian matrix ($\nabla^2 f$) is positive semi-definite). The sufficient conditions are as follows $\nabla^2 f(\mathbf{x}^*) > 0$, i.e., the Hessian matrix ($\nabla^2 f$) is positive definite. The gradient of the scalar function, a vector of size $[L \times 1]$, when differentiated by the optimization variables is defined by the expression:

$$\nabla f(\mathbf{x}) = \Gamma^T \mathbf{r}, \quad \text{(A2)}$$

where $\Gamma = [\mathbf{J_0}, \dots \ , \mathbf{J}_l, \dots \ \mathbf{J}_{L-1}] = [\frac{\partial \mathbf{r}}{\partial x_0}, \ \dots, \frac{\partial \mathbf{r}}{\partial x_l}, \dots \ \frac{\partial \mathbf{r}}{\partial x_{L-1}}]$ is a matrix of size $[K \times L]$ of first derivatives, $\mathbf{J}_l = \frac{\partial \mathbf{r}}{\partial x_l} = [\frac{\partial r_1}{\partial x_l}, \dots, \frac{\partial r_k}{\partial x_l}, \dots \frac{\partial r_K}{\partial x_l}]^T$ is a column vectors of size $[K \times 1]$, thus the Jacobian matrix element is $[\Gamma]_{kl} = \frac{\partial r_k}{\partial x_l}$. According to (A1): $\Gamma = -\Gamma_M$, where Γ_M is the matrix of first derivatives of the approximating function vector $\mathbf{M}(x)$:

$$\Gamma_M = [\frac{\partial \mathbf{M}}{\partial x_0}, \ \dots, \frac{\partial \mathbf{M}}{\partial x_l}, \dots \ \frac{\partial \mathbf{M}}{\partial x_{L-1}}]. \quad \text{(A3)}$$

In expression (A3) $\frac{\partial \mathbf{M}}{\partial x_l} = [\frac{\partial M(\mathbf{x}, \lambda_{j_a})}{\partial x_l}, \dots, \frac{\partial M(\mathbf{x}, \lambda_j)}{\partial x_l}, \dots \frac{\partial M(\mathbf{x}, \lambda_{j_b})}{\partial x_l}]^T$.

For second-order optimization methods, it is preferable to have an analytical expression for the Hessian matrix $H = \nabla^2 f(x)$. This is a matrix of size $[L \times L]$, whose element of p-th row and q-th column is described by the expression: $[H]_{(p=0:L-1)(q=0:L-1)} = \frac{\partial^2 f(\mathbf{x})}{\partial x_p \, \partial x_q}$.

The Hessian matrix of the least squares problem has the form:

$$H = \Gamma_M{}^T \Gamma_M - \sum_{k=1}^{K} H_{Mk} = \ H_a - \sum_{k=1}^{K} H_{Mk} = H_{\mathbf{a}} - H_M. \quad \text{(A4)}$$

The matrix element H_{Mk} in expression (A4) represents matrices, whose elements are given by the expressions: $[H_{Mk}]_{pq} = r_k \frac{\partial^2 M(\mathbf{x}, \lambda_k)}{\partial x_p \, \partial x_q}$. The optimization is an iterative process of successive approximations of the local minimum. The vector of variables on $(t+1)$-th iteration is expressed through the one on the previous iteration: $\mathbf{x}_{(t+1)} = \mathbf{x}_{(t)} + \mathbf{p}_{(t)}$. According to Newton's method, the vector of an iteration step

$\mathbf{p}_{(t)}$ is the solution of the system of equations: $\nabla^2 f(\mathbf{x}_{(t)})\mathbf{p}_{(t)} = -\nabla f(\mathbf{x}_{(t)})$, that, considering introduced notations (A2) and (A4), gives a system of linear equations

$$(H_a - H_M)\mathbf{p}_{(t)} = \Gamma_M{}^T\mathbf{r}. \quad \text{(A5)}$$

The Hessian and Jacobian matrices in expression (A5) are calculated for the t'th step optimization variables vector ($\mathbf{x}_{(t)}$). If we neglect second derivatives in the expression for the Hessian matrix (i.e. if we leave only first-order derivatives cross-products matrix H_a), we obtain the Gauss-Newton method. The calculation of H_a is also required when using the Levenberg-Marquart method [30]. In general case the polynomial approximation can be introduced as $V\mathbf{a}$, where $\mathbf{a}$ is the vector of coefficients and V is the Vandermonde matrix of polynomial basis functions values (some combination of independent variable exponents). For standard polynomial basis this matrix has the following form:

$$V = \begin{pmatrix} \lambda_1{}^0 & \dots & \lambda_1{}^l & \dots & \lambda_1{}^{L-2} \\ \vdots & & \vdots & & \vdots \\ \lambda_k{}^0 & & \lambda_k{}^l & & \lambda_k{}^{L-2} \\ \vdots & & \vdots & & \vdots \\ \lambda_K{}^0 & \dots & \lambda_K{}^l & \dots & \lambda_K{}^{L-2} \end{pmatrix}. \quad \text{(A6)}$$

A superscript in (A6) denotes the power exponent. It should be mentioned that the subsequent analysis is based on the matrix form of polynomial approximation and it does not require specifying the Vandermonde matrix explicitly, thus the resulting formulas can be applied to another set of polynomial basis functions (e.g. Chebyshev polynomials) by replacing the V-matrix or, more generally, any linear model of emissivity approximation. We also introduce the operation of diagonalization of a vector: let $\boldsymbol{\alpha} = (\alpha_1 \dots \alpha_k \dots \alpha_K)^T$ be a vector of size $[K \times 1]$, then its representation as a diagonal matrix of size $[K \times K]$ will be denoted as $\boldsymbol{\alpha}^D$, where $\boldsymbol{\alpha}$ vector elements stand on the main diagonal and all other elements are equal to zero. Then, taking into account (A5, A6), the vectorized expression for the approximating function of size $[K \times 1]$ in expression (8) can be put as follows:

$$\mathbf{M}(\mathbf{x}) = \mathbf{I}_c^D(x_{L-1})V\mathbf{a}, \quad \text{(A7)}$$

In (A7) $\mathbf{I}_c^D$ is a diagonal matrix of the Planck function values column-vector (corrected for the incident radiation flux). In expression (A7), the matrix V does not depend on the optimization parameters, so it does not change from iteration to iteration, that helps to reduce the computational resources of the optimization algorithm by filling V only at the start of the optimization process. Furthermore, the elements of the matrix $\mathbf{I}_c^D(x_{L-1})$ depend on only one of the optimization parameters ($x_{L-1} = T_{true}$), that simplifies the calculation of derivatives. Given expression (A3), it is convenient to represent the Jacobian matrix in a block form:

$$\Gamma_M = [\Gamma_{M1}, \mathbf{\Gamma}_{M2}], \quad \text{(A8)}$$

where Γ_{M1} is a matrix of size $[K \times (L-1)]$, the first derivatives of the target function by the parameters of emissivity approximation, and $\mathbf{\Gamma}_{M2}$ is a vector of size $[K \times 1]$, of the first derivatives with respect to temperature. Taking into account (A7), we have:

$$\Gamma_{M1} = \mathbf{I}_c{}^D V, \quad \text{(A9a)}$$

$$\mathbf{\Gamma}_{M2} = \frac{d\mathbf{I}_c^D}{dT} V\mathbf{a} = \mathbf{I}_c'^D V\mathbf{a}. \quad \text{(A9b)}$$

In expression (A9b), $\mathbf{I}_c'^D$ is the diagonalised vector of the Planck function derivatives with respect to temperature. Next, the component of the Hessian matrix, including only first-order derivatives (H_a) from (A8, A9) has the form:

$$H_a = \Gamma_M{}^T \Gamma_M = \begin{bmatrix} \Gamma_{M1}{}^T \\ \mathbf{\Gamma}_{M2}{}^T \end{bmatrix} [\Gamma_{M1}, \mathbf{\Gamma}_{M2}] = \begin{pmatrix} \Gamma_{M1}{}^T \Gamma_{M1} & \Gamma_{M1}{}^T \mathbf{\Gamma}_{M2} \\ \mathbf{\Gamma}_{M2}{}^T \Gamma_{M1} & \mathbf{\Gamma}_{M2}{}^T \mathbf{\Gamma}_{M2} \end{pmatrix} = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{b}^T & h_a \end{pmatrix} \quad \text{(A10)}$$

Thus, the matrix H_a is a block-matrix consisting of a square positive definite matrix A (see expression A11b), a vector column **b**, and scalar h_a, where, according to (A9 a, b):

$$\mathbf{b} = V^T \mathbf{I}_c^D \mathbf{I}_c'^D V\mathbf{a}, \quad \text{(A11a)}$$

$$A = V^T (\mathbf{I}_c^D)^2 V, \quad \text{(A11b)}$$

$$h_a = \mathbf{a}^T V^T (\mathbf{I}_c'^D)^2 V\mathbf{a}. \quad \text{(A11c)}$$

Now we need to obtain an exact expression for the Hessian matrix, which includes the second derivatives, $H_M = \sum_{k=1}^{K} H_{Mk}$ from eq. (A4). Most elements of the matrix H_{Mk} are equal to zero since they are the second derivatives of the coefficients of the

emissivity approximation. This matrix is symmetric and can be represented in a similar to (A10) block form:

$$\mathrm{H_{Mk}} = \begin{pmatrix} \mathrm{O} & \mathbf{H}_{\mathrm{Mk}} \\ \mathbf{H}_{\mathrm{Mk}}{}^{\mathrm{T}} & \mathrm{h_{Mk}} \end{pmatrix} \tag{A12}$$

In expression (A12), O is an all-zeros matrix of size $[(\mathrm{L}-1)\times(\mathrm{L}-1)]$, while $\mathbf{H}_{\mathrm{Mk}}$ is a column vector of size $[(\mathrm{L}-1)\times 1]$:

$$\mathbf{H}_{\mathrm{Mk}} = [\mathrm{r_k}\frac{\mathrm{dI_{ck}}}{\mathrm{dT}}\lambda_{\mathrm{k}}{}^{0}, \ldots, \mathrm{r_k}\frac{\mathrm{dI_{ck}}}{\mathrm{dT}}\lambda_{\mathrm{k}}{}^{\mathrm{l}}, \ldots, \mathrm{r_k}\frac{\mathrm{dI_{ck}}}{\mathrm{dT}}\lambda_{\mathrm{k}}{}^{\mathrm{L-2}}]^{\mathrm{T}} = \ [\mathrm{V^T}]_{\mathrm{k}}\mathrm{r_k}\frac{\mathrm{dI_{ck}}}{\mathrm{dT}}. \tag{A13}$$

In (A13), $[\mathrm{V^T}]_{\mathrm{k}}$denotes the k-th row of the Vandermonde matrix V. In (A12), $\mathrm{h_{Mk}}$ is a scalar equal to:

$$\mathrm{h_{Mk}} = \frac{\mathrm{d^2I_{ck}}}{\mathrm{dT^2}}\mathrm{r_k}\sum_{\mathrm{l}=0}^{\mathrm{L-2}}\lambda_{\mathrm{k}}{}^{\mathrm{l}}\,\mathrm{x_l}. \tag{A14}$$

Now the whole $\mathrm{H_M}$ matrix is:

$$\mathrm{H_M} = \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathrm{H_{Mk}} = \begin{pmatrix} \mathrm{O} & \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathbf{H}_{\mathrm{Mk}} \\ \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathbf{H}_{\mathrm{Mk}}{}^{\mathrm{T}} & \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathrm{h_{Mk}} \end{pmatrix} = \begin{pmatrix} \mathrm{O} & \mathbf{H}_{\mathrm{M}} \\ \mathbf{H}_{\mathrm{M}}{}^{\mathrm{T}} & \mathrm{h_M} \end{pmatrix} \tag{A15}$$

In accordance with (A13) and (A14), the following notations are introduced in (A15):

$$\mathbf{H}_{\mathrm{M}} = \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathbf{H}_{\mathrm{Mk}} = \sum_{\mathrm{k}=1}^{\mathrm{K}}[\mathrm{V^T}]_{\mathrm{k}}\mathrm{r_k}\frac{\mathrm{dI_{ck}}}{\mathrm{dT}} = \mathrm{V^T}\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}}\mathbf{r} \tag{A16}$$

$$\mathrm{h_M} = \sum_{\mathrm{k}=1}^{\mathrm{K}}\mathrm{h_{Mk}} = \sum_{\mathrm{k}=1}^{\mathrm{K}}\frac{\mathrm{d^2I_{ck}}}{\mathrm{dT^2}}\mathrm{r_k}\sum_{\mathrm{l}=0}^{\mathrm{L-2}}\mathrm{t_k}{}^{\mathrm{l}}\,\mathrm{x_l} = \mathbf{r}^{\mathrm{T}}\mathbf{I}_{\mathrm{c}}^{\prime\prime\mathrm{D}}\mathrm{V}\mathbf{a} \tag{A17}$$

In expression (A17), $\mathbf{I}_{\mathrm{c}}^{\prime\prime\mathrm{D}}$ is a diagonal matrix of size $[\mathrm{K}\times\mathrm{K}]$ of the Planck function second derivatives. Finally, the whole Hessian matrix (including both first and second derivatives) takes the following form:

$$\mathrm{H} = \mathrm{H_a} - \mathrm{H_M} = \begin{pmatrix} \mathrm{A} & \mathbf{b}-\mathbf{H}_{\mathrm{M}} \\ \mathbf{b}^{\mathrm{T}}-\mathbf{H}_{\mathrm{M}}{}^{\mathrm{T}} & \mathrm{h_a}-\mathrm{h_M} \end{pmatrix} = \begin{pmatrix} \mathrm{A} & \boldsymbol{\beta} \\ \boldsymbol{\beta}^{\mathrm{T}} & \mathrm{h} \end{pmatrix}. \tag{A18}$$

In expression (A18), taking into account (A11a-c), the following notations are introduced:

$$\boldsymbol{\beta} = \mathbf{b}-\mathbf{H}_{\mathrm{M}} = \mathrm{V^T}\mathbf{I}_{\mathrm{c}}^{\mathrm{D}}\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}}\mathrm{V}\mathbf{a} - \mathrm{V^T}\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}}\mathbf{r} = \mathrm{V^T}\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}}(\mathbf{I}_{\mathrm{c}}^{\mathrm{D}}\mathrm{V}\mathbf{a}-\mathbf{r}) \tag{A19a}$$

$$\mathrm{h} = \mathrm{h_a}-\mathrm{h_M} = \mathbf{a}^{\mathrm{T}}\mathrm{V}^{\mathbf{T}}(\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}})^2\mathrm{V}\mathbf{a} - \mathbf{r}^{\mathrm{T}}\mathbf{I}_{\mathrm{c}}^{\prime\prime\mathrm{D}}\mathrm{V}\mathbf{a} = [\mathbf{a}^{\mathrm{T}}\mathrm{V}^{\mathbf{T}}(\mathbf{I}_{\mathrm{c}}^{\prime\mathrm{D}})^2 - \mathbf{r}^{\mathrm{T}}\mathbf{I}_{\mathrm{c}}^{\prime\prime\mathrm{D}}]\mathrm{V}\mathbf{a} \tag{A19b}$$

In (A18), the matrix A is defined by expression (A11b). The resulting matrix formulas (A11a-c, A18, A19) are computationally efficient. At each iteration of the optimization algorithm, only four vectors are to be updated: $\mathbf{a}$, $\mathbf{I}_c$, $\mathbf{I}'_c$ and $\mathbf{I}''_c$, and

the Hessian and Jacobian matrices are introduced by a series of matrix multiplications. In the numerical implementation of the algorithm, it is convenient to use simple analytical formulas for the derivatives of the Planck function with respect to temperature:

$\frac{dI_{bb}}{dT} = \xi I_{bb}$, $\frac{d^2 I_{bb}}{dT^2} = \frac{d\xi}{dT} I_{bb} + \xi^2 I_{bb} = \left(2\xi^2 - \xi(\frac{2\lambda T + C_2}{\lambda T^2})\right) I_{bb}$,

where $I_{bb}(\lambda, T) = \frac{1}{\lambda^5} \frac{C_1}{\exp\exp\left(\frac{C_2}{\lambda T}\right) - 1}$, $\xi = \frac{C_2}{\lambda T^2} \frac{\exp\left(\frac{C_2}{\lambda T}\right)}{\exp\exp\left(\frac{C_2}{\lambda T}\right) - 1}$,

wavelength in μm, temperature in Kelvin, C_1=1.191043E8, $W \cdot \mu m^4/(m^2 \cdot Str)$; C_2 =14387.752, $\mu m \cdot K$.