

TP Méthodes Numériques de Base

Identification de Conductivité

Résumé et Informations Pratiques

Les méthodes numériques sont souvent développées avec l'objectif de résoudre un problème issu d'une autre discipline comme la physique, l'ingénierie ou la finance, etc. Il est donc aussi important d'appliquer les outils mathématiques acquis dans le cours que de comprendre leur théories. L'objectif de ce TP est de mettre en pratique, dans un seul sujet, des démarches variées issues des méthodes numériques introduites dans le cours.

Les consignes ci-dessous sont à respecter lors de votre réalisation de ce TP:

- **Travail en binôme.**
- Utiliser **Scilab** pour écrire vos programmes de calcul.
- Rédiger un **Compte-rendu**. Vous devez répondre à toutes les questions de l'énoncé, expliciter les méthodes employées, présenter et commenter les résultats obtenus. Il n'y a qu'un seul rapport à rendre par binôme. La qualité de la rédaction, de la synthèse, de l'analyse des résultats obtenus sont des critères importants pour la note. Le compte-rendu sera dactylographié. Nous conseillons fortement d'utiliser le logiciel **L^AT_EX**, qui est un outil extrêmement utilisé pour la rédaction d'articles scientifiques. Ce compte-rendu n'excèdera pas 10 pages et ne comportera pas de programmes. Les programmes Scilab nous seront envoyés à côté; la lisibilité du code et la pertinence des commentaires seront pris en compte dans la note du TP.
- **Remise du rapport et des programmes.** Le TP est à rendre au plus tard le **02 Mai 2017**. Il faudra déposer sur **TEIDE** votre rapport sous la forme d'un fichier .pdf ainsi que vos programmes Scilab.
- **Séances de TP et contacts.** Une séance de TP pour chaque groupe est prévue pendant le mois d'avril. Nous répondrons à vos questions sur ce TP. En dehors de ce créneau, vous pouvez vous adresser à

- chun-hsiang.tsou@univ-grenoble-alpes.fr
- arnaud.sengers@univ-grenoble-alpes.fr

Évitez de faire le TP au dernier moment.

1 Description du problème

On considère une barre métallique assimilée à un milieu unidimensionnel qui présente un défaut, réduisant localement sa conductivité thermique (voir Figure 1). La position x_d de ce défaut est inconnue et le but principal de ce TP est

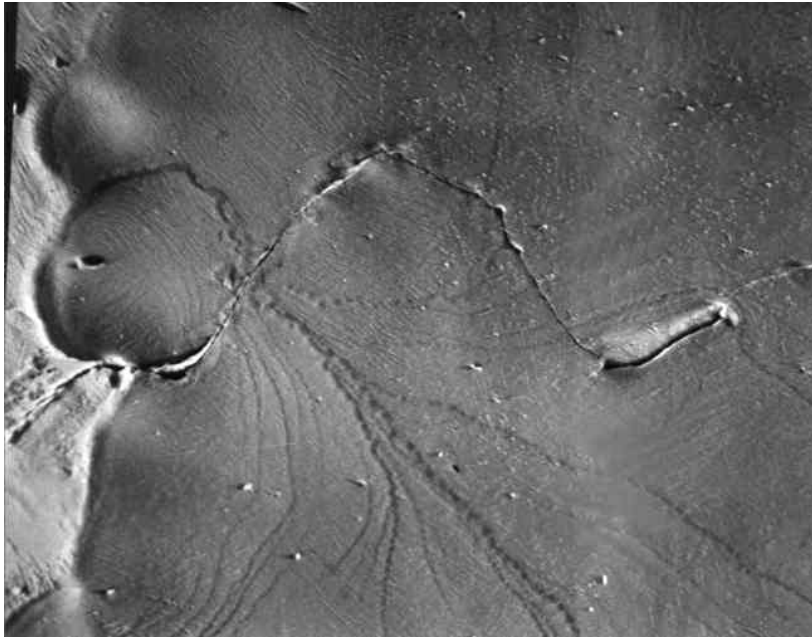


Figure 1: Photo représente la surface d'un métal, avec sur la droite un interstice créé par une fissure interne. Le défaut à détecter peut correspondre par exemple à une telle fissure interne.

de la déterminer. A cet effet, on introduit la fonction $C(x)$ qui modélise la conductivité de la barre, donnée par la formule suivante :

$$C(x) = 1 - a \exp\left(-\frac{(x - x_d)^2}{4}\right) \quad x \in [-l, l], \quad (1)$$

où $l > 0$ désigne la demi-longueur de la barre, $a \in]0, 1[$ un paramètre caractéristique du défaut métallique et $x_d \in]-l, l[$ la position du défaut dans la barre.

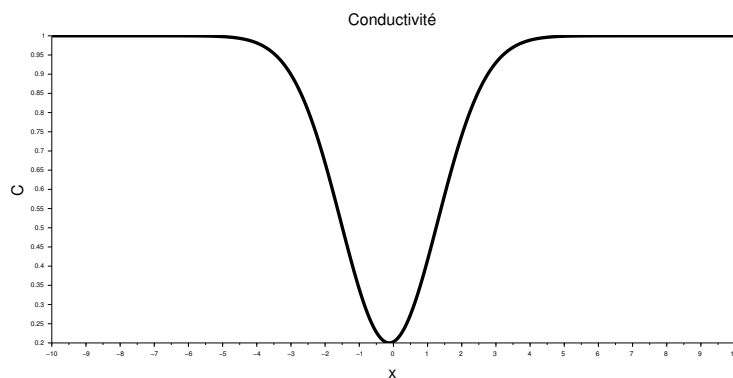


Figure 2: Conductivité - Modèle d'un défaut métallique

On désigne par $u(x, t)$ la température au point x de la barre et à l'instant t . La température vérifie l'équation de la chaleur suivante:

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}[C(x)\frac{\partial u}{\partial x}] = 0 & x \in]-l, l[, t \in [0, T], \\ u(x, 0) = u^{(0)}(x) & x \in]-l, l[, \\ u(-l, t) = u_0(t), u(l, t) = 0 & t \in [0, T], \end{cases} \quad (2)$$

avec les fonctions $u^{(0)}, u_0$ connues qu'on explicitera après.

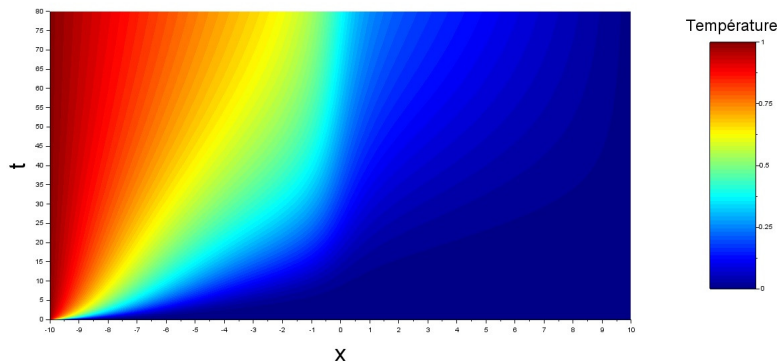


Figure 3: Profil de l'évolution de la température au cours du temps. On fixe ici $u_0(t) = 1$, ce qui correspond au cas qu'on va étudier dans les sections 4 et 5.

On donnera une mesure sur le flux de la chaleur à une extrémité de la barre,

$$F(t) := C(-l)\frac{\partial u}{\partial x}(-l, t). \quad (3)$$

L'objectif de ce TP est de reconstruire la conductivité $C(x)$, i.e. l'inconnue x_d à partir de la mesure du flux de chaleur. On appelle un tel problème un **problème inverse** sur l'équation de la chaleur.

Le TP comportera les études suivantes :

- Résolution numérique de l'équation de la chaleur par une méthode de différences finies.
- Résolution d'un système linéaire issu du schéma numérique.
- Étude de la convergence de la solution de l'équation (2) vers une solution stationnaire.
- Deux algorithmes d'optimisation.

2 La méthode des différences finies

La méthode des différences finies est la méthode la plus pratique pour traiter les équations aux dérivées partielles en dimension 1. Quand on travaille en dimension supérieure, sauf dans certains cas de géométries très particulières, cette méthode devient rapidement inutilisable. L'idée centrale de la méthode

est d'approcher les dérivées par des "différences finies", obtenues par exemple avec une formule de Taylor.

Considérant la solution u de l'équation (2), que l'on suppose de classe \mathcal{C}^2 , les développements de Taylor donnent :

$$u(x, t + h) = u(x, t) + h \frac{\partial u}{\partial t}(x, t) + \frac{h^2}{2} \frac{\partial^2 u}{\partial t^2}(x, t) + o(h^2) \quad (4)$$

$$u(x, t - h) = u(x, t) - h \frac{\partial u}{\partial t}(x, t) + \frac{h^2}{2} \frac{\partial^2 u}{\partial t^2}(x, t) + o(h^2). \quad (5)$$

Cela permet d'établir deux formules d'approximation de u selon la variable t :

Formule décentrée à droite :

$$\frac{\partial u}{\partial t}(x, t) \approx \frac{u(x, t + h) - u(x, t)}{h}, \quad (6)$$

Formule décentrée à gauche :

$$\frac{\partial u}{\partial t}(x, t) \approx \frac{u(x, t) - u(x, t - h)}{h}. \quad (7)$$

Ainsi que la formule d'approximation de la dérivée seconde selon la variable x :

Formule centrée du deuxième d'ordre :

$$\begin{aligned} \frac{\partial}{\partial x} \left[C(x) \frac{\partial u}{\partial x} \right] (x, t) \approx \\ \frac{C(x + \frac{h}{2})u(x + h, t) - (C(x + \frac{h}{2}) + C(x - \frac{h}{2}))u(x, t) + C(x - \frac{h}{2})u(x - h, t)}{h^2}. \end{aligned} \quad (8)$$

On discrétise par la suite le domaine d'étude $] -l, l[\times [0, T]$. On choisit un pas d'espace $\delta_x = \frac{2l}{n+1}$ et un pas de temps δ_t , on a,

$$-l = x_0 < x_1 < \dots < x_i = i\delta_x < x_{i+1} < \dots < x_n < x_{n+1} = l$$

et

$$0 = t_0 < t_1 < \dots < t_k = k\delta_t < t_{k+1} < \dots < t_{n_t} = T.$$

On désigne par $u_i^{(k)}$ une valeur approchée de la solution u au point (x_i, t_k) : $u_i^{(k)} \approx u(x_i, t_k)$, et on notera $C_{i\pm 1/2}$ la valeur de conductivité $C(x_i \pm \frac{\delta_x}{2})$. On utilise ensuite la formule (8) pour la dérivée seconde pour obtenir une expression de la dérivée du flux de chaleur :

$$\frac{\partial}{\partial x} \left[C(x) \frac{\partial u}{\partial x} \right] (x_i, t_k) \approx \frac{C_{i+1/2}u_{i+1}^{(k)} - (C_{i+1/2} + C_{i-1/2})u_i^{(k)} + C_{i-1/2}u_{i-1}^{(k)}}{\delta_x^2}. \quad (9)$$

Pour la dérivée temporelle, on peut utiliser au choix (6) ou (7). La façon plus générale est de combiner les deux formules avec un poids $\theta \in [0, 1]$, pour obtenir ce qu'on appelle un **θ -schéma**.

Un θ -schéma consiste en la relation suivante entre les coefficients : $\forall i = 1 \dots n, k \geq 0$,

$$\begin{aligned} \frac{u_i^{(k+1)} - u_i^{(k)}}{\delta_t} = \frac{\theta}{\delta_x^2} [C_{i+1/2}u_{i+1}^{(k+1)} - (C_{i+1/2} + C_{i-1/2})u_i^{(k+1)} + C_{i-1/2}u_{i-1}^{(k+1)}] \\ + \frac{1-\theta}{\delta_x^2} [C_{i+1/2}u_{i+1}^{(k)} - (C_{i+1/2} + C_{i-1/2})u_i^{(k)} + C_{i-1/2}u_{i-1}^{(k)}]. \end{aligned} \quad (10)$$

- Si $\theta = 0$, on dit que le **schéma** est **explicite**. On obtient alors les coefficients à t_{k+1} directement à partir de ceux à t_k sans calcul. Néanmoins, ce choix entraîne des restrictions sur le pas de temps pour que la méthode soit stable.
- Si $\theta = 1$, on dit que le **schéma** est **implicite**. Le schéma est inconditionnellement stable mais, pour passer de t_k à t_{k+1} , il est nécessaire de résoudre un système linéaire.
- Si $\theta = \frac{1}{2}$, on parle d'un **schéma** de **Crank-Nicholson**. Comme le schéma implicite, il est inconditionnellement stable, mais présente en plus l'avantage d'être plus précis que le schéma implicite avec une précision temporelle sur l'erreur en $\mathcal{O}(\delta t^2)$ (contre $\mathcal{O}(\delta t)$ pour le schéma implicite).

Dans la section 6, on choisira $\theta = \frac{1}{2}$ et la méthode de **Crank-Nicholson** pour ces raisons.

Question 1. On définit les vecteurs $U^{(k)} := (u_1^{(k)}, u_2^{(k)}, \dots, u_n^{(k)})^T$.
Montrer que le schéma (10) s'écrit de façon matricielle

$$(I + \theta \mu A)U^{(k+1)} = (I + (\theta - 1)\mu A)U^{(k)} + \mu B^{(k)}, \quad (11)$$

où $\mu = \frac{\delta_t}{\delta_x^2}$, $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice symétrique, tridiagonale qu'on explicitera, et $B^{(k)} \in \mathbb{R}^n$ avec

$$\begin{aligned} B_1^{(k)} &= C_{1/2}(u_0(t_{k+1})\theta + u_0(t_k)(1 - \theta)) \\ B_i^{(k)} &= 0 \quad \forall i = 2, \dots, n. \end{aligned}$$

Question 2. Montrer que A est définie-positive.

3 Factorisation de Cholesky dans le cas tridiagonal

Comme l'équation (11) consiste en un système linéaire $MX = Y$ avec $M = I + \theta \mu A \in \mathcal{M}_n(\mathbb{R})$, il faut choisir une méthode pour le résoudre numériquement. Comme la matrice M est tridiagonale et qu'elle est symétrique définie positive, la méthode de **factorisation de Cholesky** est particulièrement adaptée :

Théorème 3.1 (Cholesky). Soit M une matrice symétrique, définie positive, alors il existe une unique matrice triangulaire inférieure L telle que $M = LL^T$ et $L_{ii} > 0$ pour tout $i \in [1; n]$.

Une fois cette factorisation obtenue la résolution du système $MX = Y$ est rapide, et on ne calcule la factorisation qu'une fois pour toutes les itérations du θ -schéma. En effet, on a :

$$LL^T X = Y \Leftrightarrow \begin{cases} LZ = Y \\ L^T X = Z. \end{cases} \quad (12)$$

Comme la matrice A est tridiagonale, seuls les coefficients diagonaux et sous diagonaux de L sont non nuls. Ainsi, pour gagner en efficacité, on stockera la matrice L dans deux vecteurs (un de taille n pour les coefficients diagonaux et un de taille $n - 1$ pour les coefficients sous-diagonaux).

Question 3. Écrire une fonction *factorise* qui prend en argument la diagonale et la sous-diagonale d'une matrice M symétrique, définie positive et tridiagonale et qui renvoie deux vecteurs *linf* et *ldiag* contenant la sous-diagonale et la diagonale de la matrice L .

Question 4. Écrire une fonction *descente* qui calcule la solution Z du système $LZ = Y$. Cette fonction prendra pour arguments les coefficients non nuls de L (*linf*, *ldiag*) et le vecteur Y .

Question 5. Écrire une fonction *remonte* qui calcule la solution X du système $L^T X = Z$. Cette fonction prendra pour arguments les coefficients non nuls de L (*linf*, *ldiag*) et le vecteur Z .

Tester les fonctions *factorise*, *descente*, *remonte* sur un système linéaire de votre choix dont la solution est connue.

4 Problème stationnaire

On étudie maintenant le problème stationnaire, c'est-à-dire l'équation différentielle suivante :

$$\begin{cases} \frac{\partial}{\partial x} [C(x) \frac{\partial u}{\partial x}] = 0 & x \in]-l, l[\\ u(-l) = u_0 \\ u(l) = 0. \end{cases} \quad (13)$$

où u_0 est constante.

Question 6. On discrétise le problème stationnaire (13) par la méthode des différences finies.

Montrer que le problème discrétisé s'écrit sous forme matricielle $AU = B$ (où A est la matrice de la **Question 1**) et qu'il admet une unique solution.

Question 7. On prend par exemple $C(x) = \exp(-x/l)$, $u_0 = 1$.

Calculer la solution explicite de l'équation (13) et la résoudre numériquement par la méthode de différences finies, en utilisant la méthode de Cholesky implémentée dans les **questions 3-5** pour la résolution du système linéaire.

Comparer la solution explicite et celle numérique avec plusieurs $\delta_x \rightarrow 0$. (On pourra tracer les graphes de ces solutions, et calculer la norme $\|\cdot\|_\infty$ de la différence.)

5 Évolution d'une donnée initiale

On considère dans cette section que les conditions de Dirichlet aux deux bouts de la barre sont des fonctions constantes, c'est-à-dire $u_0(t) = u_0$ et on utilise la méthode de **Crank-Nicholson** (en fixant $\theta = \frac{1}{2}$). La relation (11) s'écrit donc de la manière suivante:

$$MU^{(k+1)} = NU^{(k)} + B, \quad (14)$$

où $M = I + \theta\mu A$, $N = I + (\theta - 1)\mu A$.

Question 8. Montrer que le rayon spectral de $M^{-1}N$ est strictement inférieur à 1.

Indication : on pourra commencer par vérifier que $M^{-1}N$ et A ont même

vecteurs propres et en déduire les valeurs propres de $M^{-1}N$ en fonction de celles de A .

Question 9. Dans le cas d'une donnée initiale $U^{(0)}$ quelconque, montrer que, quand $k \rightarrow \infty$, la solution discrétisée $U^{(k)}$ converge vers la solution du problème stationnaire de la section précédente.

En pratique, la résolution numérique de l'équation (2) s'effectue de manière suivante. On factorise M par Cholesky, puis pour k allant de 0 à $n_t - 1$: on calcule le second membre de (11) à partir de $U^{(k)}$ (sans stocker de grande matrice, en travaillant juste avec des vecteurs), puis $U^{(k+1)}$ par descente et remonte via les fonctions correspondantes implémentées précédemment.

Question 10. On fixe :

$$u_0(t) = 1 \quad \forall t \in [0, T], \quad u^{(0)}(x) = 0 \quad \forall x \in]-l, l[$$

Pour illustrer la convergence vers la solution stationnaire, tracer des graphes de la solution numérique en fonction de x à différents temps t_k fixés.

6 Étude du problème inverse

L'objectif premier de cette étude est de résoudre un problème inverse sur l'équation (2), c'est-à-dire reconstruire la conductivité $C(x)$ à partir de la connaissance sur le flux (3). L'étude théorique (voir [1],[2]) assure l'unicité du problème inverse, c'est-à-dire qu'il ne peut pas y avoir deux conductivités différentes qui conduisent au même flux sur les extrémités.

Théorème 6.1. Soient $C_1(x), C_2(x)$ deux fonctions régulières sur $[-l, l]$, on note u_1, u_2 les solutions de l'équation (2) associées respectivement aux conductivités $C_1(x), C_2(x)$, et F_1, F_2 leurs flux calculés par (3). Alors,

$$F_1(t) = F_2(t) \quad \forall t > 0 \Rightarrow C_1(x) = C_2(x) \quad \forall x \in [0, 1]. \quad (15)$$

Pour l'étude du problème inverse, on suppose que la conductivité satisfait la modèle (1), le problème devient donc de déterminer la position $x_d^* \in]-l, l[$. On fixe aussi les conditions de Dirichlet aux extrémités ainsi que la condition initiale:

$$u_0(t) = \left(\frac{t}{T}\right)^2 \quad \forall t \in [0, T], \quad u^{(0)}(x) = 0 \quad \forall x \in]-l, l[.$$

On échantillonne deux instants particuliers $t_{inter} := \frac{2}{3}T, t_{fin} := T$ pour représenter le flux dans les calculs qui suivent.

Pour tout calcul numérique fait dans cette section, on utilisera les paramètres suivants:

- $a = 0.8$.
- $l = 10$.
- $T = 60$.
- $n = 2000 \Rightarrow \delta_x = \frac{20}{2001}$.

- $n_t = 3000 \Rightarrow \delta_t = \frac{60}{3000}$.
- $\theta = \frac{1}{2}$.

On approche le flux (3) par la formule suivante à l'ordre 2 :

$$C(x_0)\partial_x u(x_0, t) = C\left(x_{\frac{1}{2}}\right) \frac{u(x_1, t) - u(x_0, t)}{\delta_x} - \frac{\delta_x}{2} \frac{du_0}{dt} + \mathcal{O}(\delta x^2) \quad (16)$$

Question 11. Écrire une fonction flux qui prend en argument la position du défaut x_d et qui renvoie le couple de flux $(F(t_{inter}), F(t_{fin}))$ qu'on notera $\mathcal{F}(x_d)$.

On va chercher x_d^* en minimisant une fonction de coût $J(x_d)$, qui atteint son minimum lorsque $\mathcal{F}(x_d)$ est le plus près possible du flux cible :

$$\mathcal{F}_{cible} := (-0.1, -0.18).$$

Pour toute la suite du TP, la norme $\|\cdot\|$ désignera la norme euclidienne $\|\cdot\|_2$

Question 12. Écrire et représenter graphiquement la fonction de coût J qui prend en argument x_d et renvoie la valeur

$$J(x_d) := \frac{\|\mathcal{F}(x_d) - \mathcal{F}_{cible}\|^2}{\|\mathcal{F}_{cible}\|^2}. \quad (17)$$

Pour trouver le point sur lequel le minimum de la fonction J est atteint, on présente deux algorithmes d'optimisation: **Méthode dichotomique** et **Méthode de Gauss-Newton**.

Méthode dichotomique Comme J est unimodale (elle admet x_d^* pour unique minimum et est strictement décroissante sur $[-l, x_d^*]$ et strictement croissante sur $[x_d^*, l]$), on peut appliquer une méthode de dichotomie comme suit : Pour trouver le minimum de J sur $[a, b]$ (au début $a = -l$ et $b = l$) :

- On décompose $[a, b]$ en quatre intervalles de taille égale $[a, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$, $[x_3, x_4]$ où $\delta = \frac{b-a}{4}$ et $x_i = a + i\delta$.
- En comparant les valeurs de $J(x_1), J(x_2), J(x_3)$, on peut montrer que x_d^* est soit dans $[a, x_2]$, $[x_1, x_3]$ ou $[x_2, b]$, ce qui réduit la taille de l'intervalle de recherche par 2. On recommence alors sur le nouvel intervalle jusqu'à atteindre une précision choisie.

Question 13. Expliquer comment choisir le nouvel intervalle de recherche selon les valeurs de $J(x_1), J(x_2), J(x_3)$ (une représentation graphique peut être utile). Écrire une fonction *dichotomie* qui prend en argument la fonction J et une précision ε (demi-longueur de l'intervalle final) choisie et qui renvoie x_d^* obtenu par la méthode dichotomique. On prendra $\varepsilon = 10^{-5}$.

Méthode de Gauss-Newton On revient à la définition de la fonction J (17) et on admet que la fonction flux \mathcal{F} est dérivable selon x_d . Il en est alors de même pour la fonction de coût J . Une condition nécessaire pour que x_d^* soit un minimum local de la fonction J est $J'(x_d^*) = 0$, où :

$$J'(x_d) = \frac{2\mathcal{F}'(x_d)^T(\mathcal{F}(x_d) - \mathcal{F}_{cible})}{\|\mathcal{F}_{cible}\|^2} \quad (18)$$

La condition $J'(x_d^*) = 0$ s'écrit donc :

$$\mathcal{F}'(x_d^*)^T (\mathcal{F}(x_d^*) - \mathcal{F}_{cible}) = 0. \quad (19)$$

On peut résoudre cette équation numériquement par une méthode de Newton, cela nécessite le calcul d'une dérivée de seconde ordre pour la fonction \mathcal{F} , calcul devenant compliqué. La méthode de Gauss-Newton consiste à linéariser (17) autour une approximation x_0 du minimum et calculer x_1 tel que

$$\|\mathcal{F}(x_0) - \mathcal{F}_{cible} + \mathcal{F}'(x_0)(x_1 - x_0)\| \quad \text{soit minimal} \quad (20)$$

Donc, la méthode de Gauss-Newton se résume en l'algorithme suivant:

1. Définir un test d'arrêt.
2. Choisir une estimation initiale x_0 .
3. Tant que (test d'arrêt n'est pas satisfait)
 - (a) Calculer $\mathcal{F}(x_k)$ et $\mathcal{F}'(x_k)$.
 - (b) Déterminer Δx_k tel que $\|\mathcal{F}(x_k) - \mathcal{F}_{cible} + \mathcal{F}'(x_k) \Delta x_k\|$ soit minimal.
 - (c) $x_{k+1} = x_k + \Delta x_k$.
 - (d) Incrémenter l'indice k .

Fin tant que.

Pour calculer l'avancement Δx_k dans chaque itération, on utilise la formule,

$$\Delta x_k = \frac{-\mathcal{F}'(x_k)^T (\mathcal{F}(x_k) - \mathcal{F}_{cible})}{\mathcal{F}'(x_k)^T \mathcal{F}'(x_k)}. \quad (21)$$

Question 14. Écrire un programme `minJ.sce` qui calcule le minimum x_d^* de la fonction J définie par (17) en utilisant la méthode de Gauss-Newton. On prendra le test d'arrêt $|J'(x_d^*)| \leq \varepsilon = 10^{-5}$, où $J'(x_d^*)$ est donné par (18). Pour le calcul de la dérivée numérique, on utilisera la fonction Scilab **numderivative** qui calcule une dérivée approchée de la fonction donnée en argument.

References

- [1] A.G. Ramm. An inverse problem for the heat equation. *Journal of Mathematical Analysis and Applications*, 264(2):691 – 697, 2001.
- [2] A.G. Ramm. An inverse problem for the heat equation ii. *Applicable Analysis*, 81(4):929–937, 2002.