

## Analyse du manuscrit d'André-Louis Cholesky intitulé « Sur la résolution numérique des systèmes d'équations linéaires »

par Roger Mansuy, agrégé de mathématiques, professeur de MPSI au lycée Louis-le-Grand, rédacteur en chef de la revue *Quadrature* (<http://roger.mansuy.free.fr/>)

Officier de l'artillerie, polytechnicien (X1895), André-Louis Cholesky est à la fois confronté à la réalité du terrain (et même du front où il succombera en 1918) et exigeant dans ses ambitions théoriques de formalisation. Ainsi, lorsque le devoir l'appelle à divers travaux géodésiques en France et à l'étranger, il prend le temps de rédiger un texte pour détailler la méthode (on ne parle pas encore d'algorithme) qu'il vient de mettre au point pour simplifier les calculs. C'est ce manuscrit (daté de 1910 mais resté dans les documents familiaux jusqu'en 2005) qui nous permet de redécouvrir ses travaux<sup>1</sup>. Si l'essentiel de la méthode nous était déjà parvenu, ce texte (dans une version prête pour la publication), nous éclaire sur les ambitions théoriques et pratiques de Cholesky au-delà du simple résultat.



***Figure 1. André-Louis Cholesky en 1917***

---

<sup>1</sup> Ces travaux étaient connus avant la redécouverte de ce manuscrit grâce à une note du commandant Benoît, collègue de Cholesky dans les services géographiques de l'armée : Note sur une méthode de résolution des équations normales provenant de l'application de la méthode des moindres carrés à un système d'équations linéaires en nombre inférieur à celui des inconnues (procédé du commandant Cholesky), Bulletin géodésique, 2(1924) 67-77.

Le paragraphe introductif nous rappelle la nécessité du traitement des « données expérimentales ». Dans « le cas de la compensation des réseaux géodésiques » (qui intéresse particulièrement Cholesky qui a participé à des levés en France, au Maghreb et en Roumanie...) ou d'autres « recherches de lois physiques », on est souvent ramené à résoudre numériquement de lourds systèmes linéaires dont les coefficients proviennent de mesures physiques. L'étape préliminaire consiste ici à appliquer « la méthode des moindres carrés ». Si le système admet davantage d'équations que d'inconnues (comme c'est le cas pour les levés géodésiques), il est possible (voir fréquent si l'on tient compte des erreurs de mesure) qu'il n'admette pas de solution exacte. Carl Friedrich Gauss a développé, pour retrouver la position de l'astéroïde Cérès<sup>2</sup>, une méthode efficace de résolution approchée : pour « résoudre » au mieux le système  $AX=B$  où  $A$  désigne la matrice des coefficients (mesurés) du système,  $X$  le vecteur des inconnues et  $B$  un vecteur second membre (souvent issus des mesures), il suffit de minimiser la fonction  $X \rightarrow \|AX-B\|^2$  qui, à un vecteur  $X$ , associe la somme des carrés des composantes de  $AX-B$  (cette norme euclidienne s'interprète alors comme une erreur quadratique). Ce minimum est alors obtenu pour le vecteur  $X$  solution du système d'équations dites normales (il y a cette fois-ci autant d'équations que d'inconnues)  ${}^tAAX = {}^tAB$ . On remarquera que la situation initiale se ramène ainsi à un système carré dont la matrice  $M = {}^tAA$  est symétrique (point qui sera fondamental dans l'algorithme détaillé ci-après). Cholesky ne le mentionne pas mais les matrices  $M$  ainsi obtenues sont également positives (c'est-à-dire que pour tout vecteur colonne  $X$ ,  ${}^tXMX$  est un réel positif) et en pratique la plupart sont aussi inversibles — ce qu'il utilisera sans le préciser tant c'est la règle dans les calculs géodésiques où il y a davantage d'équations linéairement indépendantes que d'inconnues.

On définit pour toute matrice  $M$ , une nouvelle matrice appelée transposée de  $M$  et notée  ${}^tM$  en intervertissant le rôle des lignes et des colonnes : les coefficients de la ligne d'indice  $i$  de  $M$  sont les coefficients de la colonne d'indice  $i$  de  ${}^tM$ . Les matrices égales à leur transposée jouissent de nombreuses propriétés : elles sont appelées « symétriques » car les

---

<sup>2</sup> En janvier 1801, l'astronome italien Giuseppe Piazzi découvre furtivement un astéroïde qu'il baptise Cérès. Malheureusement, après 41 nuits, il le perd de vue dans la lumière du jour. Gauss, supposant que Cérès admet une orbite elliptique, en cherche les paramètres par la méthode des moindres carrés et parvient ainsi à prédire l'heure et la position de sa réapparition (que Heinrich Olbers observera). Il est amusant de noter qu'après ses prouesses à l'observatoire de Göttingen, Gauss s'est lancé dans les travaux de géodésiques où il a croisé les mêmes problèmes de systèmes linéaires.

coefficients en des positions symétriques par rapport à la diagonale principale sont égaux.

@@@@@@

Le premier paragraphe véritablement mathématique détaille une remarque concernant les systèmes symétriques. Pour suivre le texte de Cholesky, notons  $T$  la matrice de coefficients  $(a_{ij}^i)_{i,j}$ ,  $Y$  le vecteur colonne de coordonnées  $(y_j)_j$ ,  $X$  celui des  $(x_i)_i$  et enfin  $C$  celui des  $(C_i)_i$ . Si l'on sait résoudre un système de matrice  $T$  (ici, le système (I) qui se réécrit  $TY+C=0$ ), alors par les mêmes opérations (aux indices près), on sait résoudre les systèmes de matrice  ${}^tT$  et donc, en particulier, le système (II)  $Y={}^tTX$ . En combinant ces deux étapes (qui en fait ne sont qu'une seule même étape de calcul), on arrive donc à la résolution du système  $T{}^tTX+C=0$  (le système (III)). De plus, la matrice  $T{}^tT$ , dont les coefficients sont notés  $A_p^q$  dans le manuscrit, présente l'avantage d'être symétrique : on a «  $A_p^q=A_q^p$  ». Ces remarques, en apparence anodines, s'avèrent être le cœur de la méthode Cholesky : si une matrice symétrique  $A$  (comme celles qui apparaissent naturellement après la minimisation des moindres carrés) s'écrit sous la forme  $A=T{}^tT$ , alors pour résoudre un système de matrice  $A$ , il suffit de savoir résoudre un système de matrice  $T$ .

*Proposons-nous donc de résoudre un système d'équations de la forme III.*

Dans cet objectif fixé par Cholesky, la méthode est désormais claire : il faut décomposer la matrice du système (III) comme un produit  $T{}^tT$  où  $T$  est une matrice d'un système (I) que l'on sait facilement résoudre. L'idée de Cholesky consiste à se placer dans « le cas où la première équation contient seulement  $y_1$ , la 2eme  $y_1, y_2$ , la 3eme  $y_1, y_2$  et  $y_3$  ainsi de suite » (p.2), c'est-à-dire dans le cas où la matrice  $T$  est triangulaire inférieure.

Pour obtenir  $T$  à partir de la matrice  $A$ , il suffit d'écrire les  $n(n+1)/2$  équations (non linéaires) obtenues en identifiant  $A$  et  $T{}^tT$ , dont les inconnues sont les coefficients de  $T$ . Cholesky remarque alors qu'il est facile de résoudre ces équations en les considérant les uns à la suite des autres :

*On voit ainsi qu'on peut calculer ligne par ligne tous les coefficients du système VI.*

En fait, il omet de signaler qu'il fait une hypothèse : pour pouvoir trouver tous les coefficients de la ligne  $p$ , il faut pouvoir diviser par  $a_p^p$  et donc que ce

coefficient soit non-nul. Ceci provient en fait du caractère inversible (ou de manière équivalente défini positif) de la matrice du système obtenu par la méthode des moindres carrés avec beaucoup d'équations indépendantes. Ceci conclut la phase de description de cette méthode qui permet la résolution efficace des systèmes linéaires symétriques.

@@@@@@

Toutefois, Cholesky ne se borne pas à donner une recette ; il ambitionne aussi d'expliquer comment mener les calculs puis de commenter en véritable numéricien (avant l'heure) l'algorithme qu'il fournit. Ainsi, le paragraphe suivant détaille comment en disposant les calculs en tableau et en utilisant convenablement une machine à calculer, on peut rendre plus fiable l'application de la méthode.

Les machines de type « Dactyle » sont des machines à calculer construites autour d'une roue dont le nombre de dents varie au cours des calculs, inventées à la fin du XIXème siècle par un ingénieur suédois. La caractéristique qui justifie la citation par le commandant Cholesky est la qualité de l'afficheur qui indique, entre autres, le signe du résultat obtenu par un code couleur simple à lire.

Après cette précision d'ordre pédagogique, Cholesky reprend le rôle du mathématicien :

*On peut mettre en évidence les avantages de cette méthode de résolution des systèmes linéaires, au point de vue de l'approximation avec laquelle les résultats sont obtenus.*

Plus précisément, il va comparer l'erreur commise en effectuant les calculs à « précision limitée » (c'est-à-dire avec un nombre fixe de décimales) dans sa méthode où  $A = T^t T$  et dans une méthode où l'on aurait décomposé  $A$  sous la forme LU avec  $L$  et  $U$  deux matrices triangulaires de coefficients  $(\beta_p^q)$  et  $(\delta_p^q)$  respectivement. En identifiant les expressions obtenues pour les coefficients de  $A$ , on vérifie que les carrés des coefficients de  $T$  dans le premier cas sont égaux à des produits de coefficients de  $L$  et de  $U$  dans le second cas, ce qui donne avec les notations du manuscrit :  $\beta_p^q \delta_p^q = (\alpha_p^q)^2$ . Une erreur  $\eta$  sur chacune des mesures entraîne, dans l'égalité précédente, des erreurs de  $(\beta_p^q + \delta_p^q) \eta = 2 \alpha_p^q \eta$ . Or le

minimum de la fonction  $(x,y) \rightarrow x+y$  pour les couples  $(x,y)$  vérifiant  $xy=a^2$  est obtenu pour  $x=y=a$ , ce que Cholesky exprime ainsi :

*On sait que le produit  $\beta_p^q \delta_p^q$  étant constant la somme de ses deux facteurs atteint son minimum lorsqu'ils sont égaux. L'erreur la plus faible que l'on puisse introduire est donc  $2 a_p^q \eta$ .*

Ainsi, Cholesky peut en déduire que sa méthode est optimale (pour la minoration de l'erreur commise à la suite des imprécisions de mesure) parmi les méthodes reposant sur la décomposition en produit de matrices triangulaires<sup>3</sup> :

*Il en résulte que le mode de résolution des systèmes linéaires qui vient d'être exposé apparaît comme celui qui fournit la meilleure approximation des calculs.*

Il détaille ensuite comment garder le bénéfice de la méthode en effectuant de manière approchée les  $n$  extractions de racines carrées en partant d'une « table de carrés » et d'un peu de pratique. La méthode proposée est attribuée à Héron d'Alexandrie, mathématicien du premier siècle de notre ère, mais Cholesky semble l'ignorer. Il explique donc comment obtenir la racine d'un nombre  $N$  en en connaissant une approximation  $n$  d'erreur  $\varepsilon$ . En effet, au premier ordre la relation  $N=(n+\varepsilon)^2$  devient  $N=n(n+2\varepsilon)$ . De là, on déduit une expression de  $\varepsilon$  puis que  $n+\varepsilon=(N/n+n)/2$  : on a remplacé l'approximation  $n$  par une approximation plus précise  $(N/n+n)/2$ . En itérant le procédé (ce qui en termes mathématiques, revient à calculer la suite récurrente définie par la fonction  $x \rightarrow (N/x+x)/2$ ), on obtient une approximation avec la précision voulue sous réserve que la première approximation soit suffisamment fiable.

@@@@@@

Toujours soucieux de la mise en pratique, il propose enfin une méthode de vérification étape par étape de la résolution du système, loin d'être superflu lorsque le temps de résolution avoisine « 4 ou 5 heures » : il s'agit d'ajouter en fin de ligne un terme opposé à la somme des coefficients de la ligne correspondante. En effectuant les opérations sur les lignes également à la dernière colonne, on doit vérifier à chaque étape que la somme sur une ligne (étendue d'un coefficient) est nulle : cet invariant des calculs permet d'éliminer

---

<sup>3</sup> Il faut toutefois mentionner que les méthodes de décomposition de type LU restent essentielles pour décomposer les matrices non symétriques. En effet, toutes les matrices carrées peuvent se décomposer sous la forme PLU où P est une matrice de permutation (donc une permutation des lignes), L une matrice triangulaire inférieure et U une matrice triangulaire supérieure.

des erreurs de résolution mais cela n'apporte rien à proprement parler à la méthode de Cholesky.

Enfin, Cholesky rappelle l'usage concret qu'il a fait de sa méthode dans ses calculs géodésiques :

*On a résolu par cette méthode plusieurs systèmes dépassant 30 équations, et en particulier un système de 56 équations. Ce dernier fait partie d'un calcul de compensation des altitudes des chaînes primordiales de la triangulation de l'Algérie.*

Cette méthode de résolution exacte est d'autant plus admirable qu'elle tranche avec les méthodes itératives développées au dix-neuvième siècle par Gauss, Seidel, Jacobi même si elle s'appuie aussi sur une minimisation préalable au sens des moindres carrées. La qualité de ce texte d'analyse numérique et le soin apporté aux explications pédagogiques font de ce texte un modèle quasiment avant-gardiste.