RAPPELS ELEMENTAIRES D'ALGEBRE MATRICIELLE ET D'ANALYSE

Laurent Desbat

TIMC-IMAG, UMR CNRS 5525,

IAB, Faculté de médecine, UJF,

38706 La Tronche.

La majorité de ce chapitre de rappels élémentaires d'algèbre et d'analyse est tirée de (Lascaux et Théodor 86) pour l'algèbre matricielle, de (Rudin 95) pour l'analyse et dans une moindre mesure pour l'algèbre, enfin de (Gasquet et Witomski 95) pour l'analyse de Fourier. Nous recommandons vivement ces trois livres français (le second est une traduction) qui sont des ouvrages de bases très solides, profonds et modernes, de calcul matriciel et d'analyse. Ils accompagneront leurs lecteurs pendant de très nombreuses années.

1 ELEMENT D'ALGEBRE LINEAIRE

1.1 Systèmes discrets :

1.1.1 Application linéaire, matrice, système d'équations linéaires.

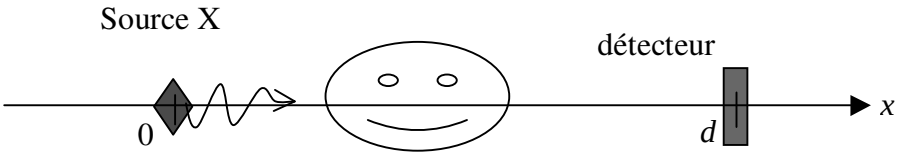
Une matrice est un tableau qui permet de représenter une application linéaire entre deux espaces vectoriels de dimensions finis. En effet, si $\mathbf{A}$ est une application linéaire de E de dimension n de base $(e_1, e_2, \cdots, e_n)$ dans F de base $(f_1, f_2, \cdots, f_m)$ de dimension m, alors notons $\mathbf{A}(e_j) = \sum_{i=1}^{m} a_{i,j} f_i$. Soit $x \in E, x = \sum_{j=1}^{n} x_j e_j$, avec x_i et $a_{i,j}$ élément de $\mathbf{C}$ ou $\mathbf{R}$, alors

$$\mathbf{A}(x) = \mathbf{A}(\sum_{j=1}^{n} x_j e_j) = \sum_{j=1}^{n} x_j \mathbf{A}(e_j) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j} x_j \right) f_i .$$

La matrice **A** représente l'application linéaire $\mathcal{A}$ sous la forme d'un tableau 2D $m \times n$ d'élément courant $a_{i,j}, i = 1,\cdots,m;\ j = 1,\cdots,n$. On remarquera que la matrice dépend du choix des bases de *E* et *F*. Comme nous allons le voir dans l'exemple ci-dessous, de très nombreux problèmes se ramènent à la résolution d'un système linéaire $\mathcal{A}x = y$, en particulier dans le domaine médical et biologique. Nous pouvons remarquer que si $y = \sum_{i=1}^{m} y_i f_i$ alors $\mathcal{A}x = y \Leftrightarrow \sum_{i=1}^{m}\left(\left(\sum_{j=1}^{n} a_{i,j}x_j\right) - y_i\right)f_i = 0$. Comme $(f_1, f_2, \cdots, f_m)$ est une base de *F* donc un système libre, ceci revient à résoudre :

$$\sum_{j=1}^{n} a_{i,j}x_j - y_i = 0, \forall i = 1,\cdots,m \Leftrightarrow \mathbf{Ax} = \mathbf{y}$$

Les transformations linéaires sont assez courantes dans la nature, même parfois cachées. Par exemple, le modèle de la formation des images en radiologie est celui de l'atténuation exponentielle d'un faisceau d'un rayon X à travers les tissus qu'il traverse suivant une droite de propagation d'axe *(O,x)*.



Si nous notons *f(x)* la fonction d'atténuation en un point *x* de la droite parcourue par le faisceau de rayons X alors, si I_0 est l'intensité du rayon X à l'origine, si I_d est l'intensité du rayon X sur le détecteur en *d*, alors nous pouvons montrer expérimentalement que :

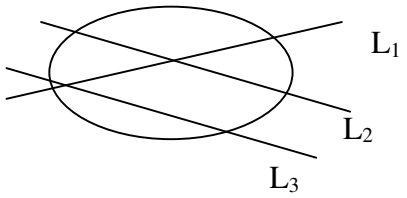
$$I_d = I_0 e^{-\int_0^d f(x)dx}.$$

Cette transformation n'est pas linéaire, mais si on passe au logarithme, alors

$$\ln\left(\frac{I_0}{I_d}\right) = \int_0^d f(x)dx$$

est bien une transformation linéaire en *f* car l'intégrale est linéaire en *f*.

Supposons que nous travaillons maintenant dans un plan de coupe du crâne et que nous pouvons disposer de plusieurs angles d'incidence pour notre système de radiographie, c'est à dire que nous pouvons mesurer *f* suivant plusieurs droites L_i du plan :

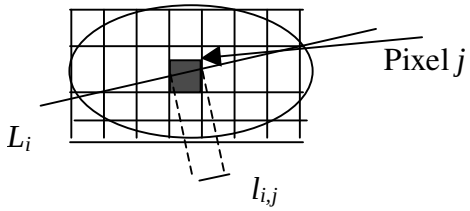


Nous avons alors accès à un ensemble de mesures : $g_i = \int_{L_i} f(x)dx, i = 1, \cdots, m,$ de la fonction *f*. Bien entendu, il n'est pas envisageable a priori, d'identifier une fonction (un point d'un ensemble de dimension infinie), à partir d'un ensemble fini de mesures. Ainsi, nous chercherons en général à discrétiser f, c'est à dire à la modéliser dans un espace de dimension finie. Ici, nous cherchons à reconstruire une image d'une coupe de densité et nous pourrons représenter la fonction *f* comme la somme d'indicatrices de pixels :

$$f(x) = \sum_{i=1}^{n} f_i e_i(x),$$

avec $e_i(x)$ indicatrice du i^{eme} pixel :

$$e_i(x) = \begin{cases} 1 & \text{si } x \text{ est dans le pixel} \\ 0 & \text{sinon} \end{cases}.$$



Le modèle de mesure devient alors

$$g_i = \int_{L_i} f(x)dx = \sum_{j=1}^{n} \left(\int_{L_i} e_j(x)dx \right) f_j = \sum_{j=1}^{n} l_{i,j} f_j, i = 1, \cdots, m,$$

avec $l_{i,j}$ est la longueur de l'intersection de la droite L_i avec le pixel *j*. Identifier *f* à partir du vecteur de mesure **g** revient alors à résoudre le système linéaire **g=Lf** où l'élément courant de

la matrice L est $l_{i,j} = \int_{L_i} e_j(x)dx$. Nous remarquons ici que $\mathbf{L}$ est une application linéaire de E dans F, avec E est l'ensemble des images en niveaux de gris sur les pixels e_i et $F = \mathbb{R}^m$.

1.1.2 Surjectivité et image, injectivité et noyau, stabilité et conditionnement, rang.

Confronté à la résolution d'un système linéaire, nous nous posons trois questions :

- *Le système a-t-il une solution ?* C'est la question de *l'existence.* La réponse à cette question est liée à l'image de $\mathbf{A}$: $\mathrm{Im}(\mathbf{A}) = \{y \in F, \exists x \in E \text{ avec } A(x) = y\}$. En effet, le problème n'aura de solution à coup sûr que si $\mathrm{Im}(\mathbf{A}) = F$. Dans ce cas nous disons que $\mathbf{A}$ est surjective. Dans le cas contraire, l'existence d'une solution est très improbable en pratique : en effet, les expérimentations dont sont issues nos données, le codage informatique des nombres (arrondis), introduisent des erreurs et le système que nous devons résoudre est $\mathbf{A}x = y = y_0 + \varepsilon$ avec $y_0 = \mathbf{A}x_0$ et l'erreur ε n'a aucune raison d'être dans Im($\mathbf{A}$).
- *La solution est-elle unique ?* C'est la question de *l'unicité.* La réponse à cette question est liée au noyau de $\mathbf{A}$: $\mathrm{Ker}(\mathbf{A}) = \{x \in E, A(x) = 0\}$ En effet, si x_1 et x_2 sont deux solutions du système linéaire :

  $$\mathbf{A}(x_1) = \mathbf{A}(x_2) \Leftrightarrow \mathbf{A}(x_1 - x_2) = 0 \Leftrightarrow (x_1 - x_2) \in \mathrm{Ker}(\mathbf{A}).$$

  Si $\mathrm{Ker}(\mathbf{A})$ est réduit au singleton $\{0\}$, alors la solution est unique, sinon, $\mathrm{Ker}(\mathbf{A})$ est un sous-espace vectoriel de E de dimension supérieure ou égale à 1 et l'ensemble des solutions est infini et peut être représenté sous la forme : $x_0 + \mathrm{Ker}(\mathbf{A})$ où x_0 est une solution du système.
- *La résolution est-elle stable ?* C'est la question de *la stabilité.* Pour aborder cette question on suppose que $\mathbf{A}$ est inversible, donc surjective (existence) et injective (unicité). On note $\mathbf{A}^{-1}$ l'inverse de $\mathbf{A}$. D'un point de vue théorique, la stabilité est liée à la continuité de $\mathbf{A}^{-1}$. En effet, si $\mathbf{A}^{-1}$ n'est pas continue,

alors nous n'avons plus l'existence d'une constante M telle que $\|\mathcal{A}^{-1}y\| \le M\|y\|$ et si une suite y_n converge vers y alors $x_n = \mathcal{A}^{-1}y_n$ ne converge pas nécessairement vers $x = \mathcal{A}^{-1}y$. D'un point de vue numérique, c'est une catastrophe car dès que y sera entaché d'erreur nous ne pourrons plus garantir que la solution est proche de x ($\mathcal{A}^{-1}(y+\varepsilon) = \mathcal{A}^{-1}y + \mathcal{A}^{-1}\varepsilon$ avec $\mathcal{A}^{-1}\varepsilon$ arbitrairement grand) et même si y est sans erreur (précision infinie?) nous ne pourrons pas garantir qu'une solution x_n, construite de manière itérative telle que $\mathcal{A}x_n = y_n$, converge vers x, même si y_n converge vers y

Ces questions se posent en dimension quelconque pour un opérateur linéaire. Bien entendu, en dimension finie, nous remplaçons $\mathcal{A}$ par la matrice **A** qui le représente. L'image de la matrice **A** est alors un sous espace vectoriel de $\mathbf{R}^m$. On appellera rang de **A**, noté rg(**A**), la dimension de $\mathrm{Im}(\mathcal{A})$. Pour toute matrice, nous avons : $\mathrm{rg}(\mathbf{A}) + \dim(\mathrm{Ker}(\mathbf{A})) = n$. Nous remarquerons que l'injectivité implique que $n \le m$ car n vecteurs en dimension $m<n$ ne peuvent pas être libres, or si $\exists\lambda_i$ non tous nuls tels que $\sum_{i=1}^{n}\lambda_i \mathbf{A}e_i = 0$, alors $\sum_{i=1}^{n}\lambda_i e_i$ est un élément non nul du noyau. La surjectivité implique que $m \le n$ car sinon l'image de **A**, de dimension inférieure ou égale à n, est nécessairement un sous espace vectoriel de dimension strictement inférieure à m. La question de la stabilité est alors plus simple. En dimension finie, tout opérateur linéaire est continu. L'analyse de la stabilité de la résolution d'un système linéaire est la suivante : si **y** est entaché de l'erreur **dy**, alors **A(x+dx)=y+dy** avec **Ax=y** et **Adx=dy**, nous avons $\|\mathbf{y}\| \le M_{\mathbf{A}}\|\mathbf{x}\|$ et $\|\mathbf{dx}\| \le M_{\mathbf{A}^{-1}}\|\mathbf{dy}\|$ soit

$$\frac{\|\mathbf{dx}\|}{\|\mathbf{x}\|} \le M_{\mathbf{A}} M_{\mathbf{A}^{-1}} \frac{\|\mathbf{dy}\|}{\|\mathbf{y}\|}$$

Nous constatons qu'une erreur relative sur **y** s'amplifie en une erreur relative sur **x** par le facteur $M_A M_{A^{-1}}$ au plus. Dans le cas de matrices diagonales $\mathbf{D} = diag(d_1, d_2, \cdots, d_n)$, avec pour simplifier $|d_1| \le |d_2| \le \cdots \le |d_n|$, la plus petite borne pour $M_{\mathbf{A}} M_{\mathbf{A}^{-1}}$ c'est à dire, le

conditionnement du problème de résolution du système linéaire est le rapport des valeurs propres extrêmes $\frac{|d_n|}{|d_1|}$. En pratique, nous travaillons avec la norme Euclidienne.

1.1.3 Matrices particulières (carrée, triangulaire, normale, symétrique, …)

Les matrices carrées sont les matrices $n \times n$. Ce sont les seules qui aient une chance d'être inversibles. Une condition nécessaire et suffisante d'inversibilité est que le déterminant de la matrice soit non nul. Cette condition n'est jamais utilisée numériquement pour des grandes matrices car l'évaluation d'un déterminant est numériquement coûteuse et potentiellement instable. De plus, le nombre réel zéro n'a pas toujours grand sens en machine après de nombreuses opérations. Deux matrices **A** et **B** sont dites semblables lorsqu'il existe une matrice inversible **S** telle que $\mathbf{A}=\mathbf{SBS}^{-1}$.

Une matrice triangulaire supérieure est une matrice **A** telle que $a_{i,j} = 0$ si i<j. Elle est dite triangulaire inférieure si $a_{i,j} = 0$ si i>j. Une matrice est diagonale si $a_{i,j} = 0$ si $i \neq j$. L'intérêt des matrices diagonales ou triangulaires est qu'elles sont facilement inversibles (lorsqu'elle sont carrées).

carrée | triangulaire inférieure | triangulaire supérieure | diagonale

La transposée de la matrice $\mathbf{A}, m \times n$ est la matrice $\mathbf{A}^t, n \times m$ telle que $a^t_{i,j} = a_{i,j}$. La matrice adjointe est la conjuguée de la transposée : $\mathbf{A}^{\#} = \overline{\mathbf{A}}^t$. Lorsque la matrice **A** est réelle, transposée et conjuguée sont confondues. Une matrice **A** telle que $\mathbf{A}=\mathbf{A}^t$ est dite symétrique. Une matrice **A** telle que $\mathbf{A}=\mathbf{A}^{\#}$ est dite hermitienne. Une matrice symétrique est hermitienne si et seulement si elle est réelle. Une matrice **A** telle que $\mathbf{A}^{-1}=\mathbf{A}^{\#}$ est dite unitaire.

Une matrice **A** telle que $\mathbf{AA}^{\#} = \mathbf{A}^{\#}\mathbf{A}$ est dite normale. Les matrices hermitiennes et unitaire sont normales. L'intérêt de la notion de matrice conjuguée réside dans le fait que si on désigne par $\langle w, x\rangle_{\mathbb{C}^n} = \sum_{i=1}^{n} w_i \overline{x}_i$ le produit scalaire sur $\mathbb{C}^n$ alors

$$\langle \mathbf{Ax}, \mathbf{y}\rangle_{\mathbf{C}^m} = \langle \mathbf{x}, \mathbf{A}^{\#}\mathbf{y}\rangle_{\mathbf{C}^n} \text{ (qui se généralise en } \langle Ax, y\rangle_F = \langle x, A^{\#}y\rangle_E)$$

Pour une matrice **A** rectangulaire nous avons le résultat fondamental suivant :

$$\mathrm{Ker}\mathbf{A}^{\#} = (\mathrm{Im}\,\mathbf{A})^{\perp} \text{ et } \mathrm{Im}\,\mathbf{A}^{\#} = (\mathrm{Ker}\mathbf{A})^{\perp}.$$

En effet, d'une part, on montre la première égalité par

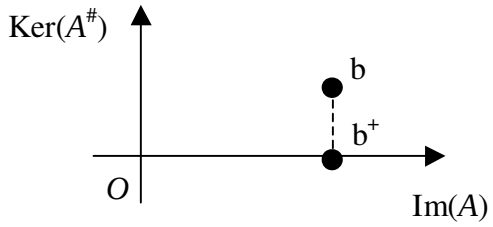
$$(\mathrm{Im}\,\mathbf{A})^{\perp} = \{\mathbf{y}, \langle \mathbf{Ax}, \mathbf{y}\rangle_{\mathbf{C}^m} = 0, \forall \mathbf{x} \in \mathbb{C}^n\} = \{\mathbf{y}, \langle \mathbf{x}, \mathbf{A}^{\#}\mathbf{y}\rangle_{\mathbf{C}^n} = 0, \forall \mathbf{x} \in \mathbb{C}^n\} = \{\mathbf{y}, \mathbf{A}^{\#}\mathbf{y} = \mathbf{0}\} = \ker \mathbf{A}^{\#}$$

La seconde s'obtient alors facilement en remplaçant dans la première **A** par $\mathbf{A}^{\#}$ et en remarquant que $(\mathbf{A}^{\#})^{\#} = \mathbf{A}$. Nous obtenons $\mathrm{Ker}\mathbf{A} = (\mathrm{Im}\,\mathbf{A}^{\#})^{\perp}$ et donc, puisqu'en dimension finie pour tout sous espace *G*, $(G^{\perp})^{\perp} = G$, $(\mathrm{Ker}\mathbf{A})^{\perp} = \mathrm{Im}\,\mathbf{A}^{\#}$. Attention, la seconde égalité ne se généralise pas systématiquement en dimension infinie pour une application linéaire A !

1.2 Résolution d'un système linéaire.

1.2.1 Solution et solution généralisée.

Nous avons vu que pour déterminer de manière unique la solution d'un système linéaire $\mathbf{Ax} = \mathbf{b}$, il faut que la matrice soit carrée et régulière, c'est à dire que son noyau soit réduit au singleton $\{0\}$. Cependant nous souhaitons donner un sens à la solution de systèmes non inversibles car ils sont fréquents. Le premier problème à résoudre est celui de l'existence qui est lié à la non-surjectivité de A, c'est à dire pour les matrices au fait que rg*(***A***)<m*. La généralisation naturelle est la notion de la solution au sens des moindres carrés. Nous chercherons donc l'ensemble des **x** qui minimisent $\|\mathbf{Ax} - \mathbf{b}\|^2$. Ceci revient simplement à projeter orthogonalement **b** sur l'image de **A** :

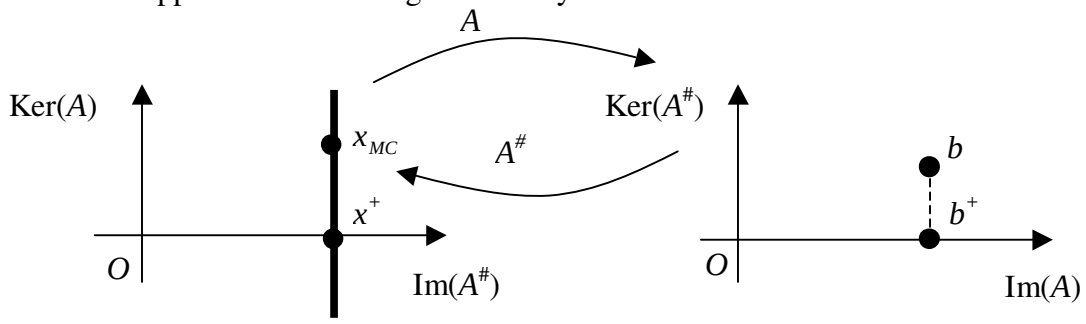


On définit alors l'unique projection $\mathbf{b}^+$ de $\mathbf{b}$ sur l'image de $\mathbf{A}$. L'ensemble des solutions de moindres carrés est alors l'ensemble des solutions à l'équation $\mathbf{Ax} = \mathbf{b}^+$. La caractérisation de $\mathbf{b}^+$ projection orthogonale de $\mathbf{b}$ sur l'image de $\mathbf{A}$ est simplement que $\mathbf{b} - \mathbf{b}^+$ est orthogonal à l'image de $\mathbf{A}$: $\forall \mathbf{y} \in \mathrm{Im}(\mathbf{A}), \left\langle \mathbf{y}, \mathbf{b} - \mathbf{b}^+ \right\rangle_{\mathbf{C}^m} = 0$ ou de manière équivalente $\forall \mathbf{x} \in \mathbf{C}^n, \left\langle \mathbf{Ax}, \mathbf{b} - \mathbf{b}^+ \right\rangle_{\mathbf{C}^m} = 0$ soit encore $\forall \mathbf{x} \in \mathbf{C}^n, \left\langle \mathbf{x}, \mathbf{A}^{\#}(\mathbf{b} - \mathbf{b}^+) \right\rangle_{\mathbf{C}^m} = 0$. Nous obtenons donc l'équation caractéristique définissant $\mathbf{b}^+$ l'unique élément de l'image de $\mathbf{A}$ telle que $\mathbf{A}^{\#}(\mathbf{b} - \mathbf{b}^+) = \mathbf{0}$. En remplaçant $\mathbf{b}^+$ par $\mathbf{Ax}$ nous obtenons le système d'équations normales associées à un problème de moindres carrés :

$$\mathbf{A}^{\#}\mathbf{Ax} = \mathbf{A}^{\#}\mathbf{b}$$

Ce système linéaire admet toujours une solution. Il est facile de voir que $\ker(\mathbf{A}) = \ker(\mathbf{A}^{\#}\mathbf{A})$. D'une part nous avons $\ker(\mathbf{A}) \subset \ker(\mathbf{A}^{\#}\mathbf{A})$ car si $\mathbf{Ax} = 0$ alors $\mathbf{A}^{\#}\mathbf{Ax} = 0$. D'autre part $\ker(\mathbf{A}^{\#}\mathbf{A}) \subset \ker(\mathbf{A})$ car si $\mathbf{A}^{\#}\mathbf{Ax} = 0$ alors nous obtenons $0 = \left\langle \mathbf{x}, \mathbf{A}^{\#}\mathbf{Ax} \right\rangle_{\mathbf{C}^n} = \left\langle \mathbf{Ax}, \mathbf{Ax} \right\rangle_{\mathbf{C}^m} = \left\| \mathbf{Ax} \right\|$ donc $\mathbf{Ax} = \mathbf{0}$. Ainsi, l'ensemble des solutions de moindres carrés est la somme d'une équation particulière et du noyau de $\mathbf{A}$. La solution particulière qui est choisie habituellement est la solution de moindres carrés et de moindre norme. Elle est appelée solution normale. La solution normale, donc de moindre carrés et de moindre norme, est la projection de l'origine sur l'ensemble des solutions moindres carrées (qui n'est autre qu'un espace affine translaté de $\ker(\mathbf{A})$). La projection orthogonale x^+ de l'origine est caractérisée par $\left\langle \mathbf{x}^+, \mathbf{x}^+ - \mathbf{x}_{MC} \right\rangle_{\mathbf{C}^n} = 0$ pour toutes les solutions de moindres carrés $\mathbf{x}_{MC}$. Or si $\mathbf{x}_{MC}$ décrit l'ensemble des solutions de moindres carrés alors $\mathbf{x}^+ - \mathbf{x}_{MC}$

décrit ker(**A**), donc $\left\langle \mathbf{x}^+, \mathbf{x}^+ - \mathbf{x}_{MC} \right\rangle_{\mathbf{C}^n} = 0$ implique que $\mathbf{x}^+ \in \left(\ker(\mathbf{A})\right)^\perp = \mathrm{Im}(\mathbf{A}^\#)$. C'est donc l'unique solution appartenant à l'orthogonal du noyau.



La notion de solution normale se généralise pour des opérateurs linéaires A en dimension infinie dès que l'image de A est fermée (c'est une condition nécessaire pour définir la projection et elle n'est pas toujours garantie).

1.2.2 Système carré inversible

Très rapidement, il existe deux grandes familles de résolution numérique des systèmes linéaires. La première est celle des méthodes dites directes. L'idée en générale est de factorisée la matrice **A** sous la forme **A=LU** de telle sorte que la résolution des systèmes linéaires associée aux matrices **U** et **L** soit aisée. Par exemple, la factorisation de Gauss connue sous le nom de méthode du pivot, cherche à obtenir **L** triangulaire inférieure et **U** triangulaire supérieure. Pour résoudre **Ax=b**, nous résolvons alors **LUx=b** en 2 étapes : résolution de **Ly=b** suivi de la résolution de **Ux=y.** L'ordre de grandeur du coût, en nombre d'opérations flottantes, de la résolution d'un système linéaire par une méthode directe est souvent le cube de la dimension de la matrice. La seconde famille est celle des méthodes itératives: une suite de vecteurs $\mathbf{x}_k$ est construite de telle manière à ce qu'elle converge vers la solution. Les méthodes de Jacobi, de Gauss-Seidel, avec relaxation SOR ou SSOR, plus généralement les méthodes de factorisation incomplète, et surtout le gradient conjugué sont les méthodes les plus célèbres. Les méthodes itératives sont souvent basées sur le produit par la matrice **A** ou des sous matrices de **A**.Elles sont donc particulièrement intéressantes si la matrice est creuse, c'est à dire essentiellement remplie de 0.

Pour plus de précision nous conseillons la lecture de l'excellent ouvrage (Lascaux et Theodor, 1995).

1.2.3 Système singulier

La résolution d'un système singulier (matrice rectangulaire, noyau non réduit au singleton $\{0\}$) est traitée généralement au sens des moindres carrés et de moindre norme. L'outil principal est la décomposition en valeurs singulières. On peut montrer que toute matrice $\mathbf{A}\, m\times n$, à coefficients complexes ou réels, peut se décomposer sous la forme $\mathbf{A} = \mathbf{UDV}$, avec $\mathbf{U}$ matrice unitaire $m\times m$, $\mathbf{V}$ matrice unitaire $n\times n$, et $\mathbf{D}$ est une matrice diagonale $m\times n$. Les éléments diagonaux de la matrice $\mathbf{D}\, m\times n$ sont simplement les racines carrées de la matrice $\mathbf{A}^{\#}\mathbf{A}$. On définit le pseudo-inverse $\mathbf{D}^{+}\, n\times m$ d'une matrice diagonale comme étant la matrice diagonale telle que

$$d_{i,i}^{+} = \begin{cases} \dfrac{1}{d_{i,i}} & \text{si} \quad \mathrm{d_{i,i}} \neq 0 \\ 0 & \text{Sinon} \end{cases}$$

La matrice pseudo-inverse est définie par $\mathbf{A}^{+} = \mathbf{V}^{\#}\mathbf{D}^{+}\mathbf{U}^{\#}$ et la solution de moindres carrés et de moindres normes est $\mathbf{x}^{+} = \mathbf{A}^{+}\mathbf{b}$.

Nous mentionnons aussi que le gradient conjugué, appliqué sur la matrice $\mathbf{A}^{\#}\mathbf{A}$ dans la résolution des équations normales, converge vers $\mathbf{x}^{+}$.

1.3 Valeurs propres, vecteurs propres

1.3.1 Définition, spectre et rayon spectral

Soit **A** une matrice carrée d'ordre *n*, on appelle valeur propre de **A** le nombre $\lambda \in \mathbb{C}$ pour lequel il existe un vecteur **u** non nul de $\mathbb{C}^n$ appelé vecteur propres tel que $\mathbf{Au} = \lambda\mathbf{u}$. Le spectre de **A**, noté Sp(**A**), est l'ensemble des valeurs propre de **A**. Le rayon spectral de **A** est le nombre $\rho(\mathbf{A}) = \max_{1\le i\le n} |\lambda_i|$.

Nous remarquons que si **u** est vecteur propre de **A** alors $\forall \alpha \in \mathbb{C}(\alpha \neq 0), \alpha\mathbf{u}$ est aussi un vecteur propre de **A**.

1.3.2 Polynôme caractéristique et multiplicité

Lorsque $\lambda \in Sp(\mathbf{A})$, alors $\exists \mathbf{u} \in \mathbf{C}^n, \mathbf{u} \neq 0, \mathbf{Au} - \lambda\mathbf{u} = 0$. Ceci signifie que :

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 .$$

On appelle polynôme caractéristique de la matrice **A**, le polynôme de degré *n* :

$$p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) .$$

Les racines du polynôme caractéristique sont les valeurs propres de **A**. Si λ est un zéro du polynôme caractéristique de multiplicité *k*, nous disons que la valeur propre λ est de multiplicité *k*. Nous pouvons remarquer que deux matrices semblables ont même polynôme caractéristique. En effet, si $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ alors :

$$\det(\mathbf{B} - \lambda\mathbf{I}) = \det(\mathbf{S}^{-1}\mathbf{AS} - \lambda\mathbf{I}) = \det(\mathbf{S}^{-1})\det(\mathbf{A} - \lambda\mathbf{I})\det(\mathbf{S}) = \det(\mathbf{A} - \lambda\mathbf{I}) .$$

Ceci montre que les valeurs propres sont caractéristiques de l'opérateur linéaire $\mathcal{A}$ et non pas de la matrice **A** qui le représente.

Soit **u** vecteur propre de **A** associé à la valeur propre λ. Nous avons les propriétés élémentaires suivantes :

- $\mathbf{A}^k\mathbf{u} = \lambda^k\mathbf{u}, \forall k \in \mathbb{Z}$, en particulier pour *k*= *-1*.
- $p(\mathbf{A})\mathbf{u} = p(\lambda)\mathbf{u}$ pour tout polynôme *p*.

1.3.3 Diagonalisation

Nous allons d'abord énoncer un théorème général puis des conditions suffisantes à la diagonalisation.

> **Théorème** Soit **A** une matrice carré d'ordre *n*. **A** est diagonalisable si et seulement si **A** possède *n* vecteurs propres $\mathbf{u}_i$ linéairement indépendants.
>
> Alors **A** peut se factoriser sous la forme $\mathbf{A} = \mathbf{SDS}^{-1}$ où **D** est la matrice diagonale formée des vecteurs propres. La $i^{ème}$ colonne de **S** est un vecteur propre $\mathbf{u}_i$ associé à la valeur propre λ_i.

En effet, si **A** possède *n* vecteurs propres $\mathbf{u}_i$ linéairement indépendants alors notons **S** la matrice dont la $i^{ème}$ colonne est le vecteur propre $\mathbf{u}_i$ $\mathbf{S} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n]$ alors **AS=SD**.

Comme **S** est inversible, $\mathbf{A} = \mathbf{SDS}^{-1}$. Réciproquement, si $\mathbf{A} = \mathbf{SDS}^{-1}$ alors les colonnes de **S** sont des vecteurs propres indépendants associés aux valeurs propres sur la diagonale de **D**.

> **Proposition** : les vecteurs propres associés à des valeurs propres distinctes sont linéairement indépendants.

Démonstration laissée en exercice (par récurrence sur le nombre de vecteurs propres). En corollaire de cette proposition, nous obtenons que si toutes les valeurs propres de **A** sont distinctes alors *A* est diagonalisable. Ce qui peut poser des problèmes, ce sont donc les valeurs propres multiples. En effet la matrice :

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

a pour polynôme caractéristique $(\lambda - 1)^2$ et donc la valeur propre 1 est double. Elle ne peut pas être diagonalisable sinon $\mathbf{A} = \mathbf{SIS}^{-1} = \mathbf{SS}^{-1}\mathbf{I} = \mathbf{I}$ **,** or clairement **A** n'est pas l'identité !
Nous admettrons les résultats fondamentaux suivants

> **Théorème de Schur** Toute matrice carrée A peut se décomposer sous la forme
>
> $$\mathbf{A} = \mathbf{UTU}^{\#}$$
>
> où **U** est une matrice unitaire et **T** est une matrice triangulaire supérieure.

> **Théorème** Une matrice **A** est normale si, et seulement si, il existe une matrice unitaire **U** telle que
>
> $$\mathbf{A} = \mathbf{UDU}^{\#}$$
>
> **D** est la matrice diagonale formée des valeurs propres de **A**. Autrement dit, une matrice normale est diagonalisable dans une base orthonormée.

Le théorème précédent est une conséquence directe du théorème de Schur (essentiel en pratique) si on montre au préalable (exercice) qu'une matrice triangulaire supérieure normale est nécessairement diagonale. Le théorème de Schur permet donc de conclure que les matrices hermitiennes, réelles symétriques ont des valeurs propres réelles et sont diagonalisables par une transformation unitaire. Les matrices unitaires sont elles même diagonalisables par transformation unitaire et ont des valeurs propres de module égal à 1.

1.4 Calcul des valeurs propres

Il n'est pas raisonnable ici encore de parcourir en seulement quelques lignes les méthodes numériques du calcul des valeurs propres d'une matrice. En résumé, ce sont en général des méthodes itératives. La méthode la plus utilisée est appelée l'algorithme QR basée sur la factorisation d'une matrice sous la forme **A=QR** où **Q** est une matrice unitaire et **R** est une matrice triangulaire supérieure (algorithme direct en n^3). Nous nous bornerons, ici, à évoquer l'idée de la méthode de la puissance. Elle permet de calculer la valeur propre de plus grand module d'une matrice lorsqu'elle est unique et le vecteur propre associé.

L'algorithme de la puissance itérée est fort simple : soit $\mathbf{x}_0 \neq \mathbf{0}$, on construit l'itération $\mathbf{x}_{k+1} = \frac{\mathbf{A}\mathbf{x}_k}{\|\mathbf{A}\mathbf{x}_k\|}$; Si on considère une matrice diagonale **D** dont la première valeur (propre) est de module strictement plus grand que les autres (**D** = diag[d_1,d_2,…,d_n], $|d_1|>|d_i|$ pour i>1) il est clair que :

$\mathbf{x}_{k+1} = \frac{\mathbf{D}\mathbf{x}_k}{\|\mathbf{D}\mathbf{x}_k\|} = \frac{\mathbf{D}^{k+1}\mathbf{x}_0}{\prod_{i=0}^{k}\|\mathbf{D}\mathbf{x}_i\|}$ aura tendance à s'aligner sur le vecteur (1,0,…,0) premier vecteur propre de **D**. En effet, pour simplifier choisissons $\mathbf{x}_0 = (1,1,\cdots,1)$, alors :

$$\frac{\mathbf{D}^{k+1}\mathbf{x}_0}{d_1^{k+1}} = \left(1, \frac{d_2^{k+1}}{d_1^{k+1}}, \cdots, \frac{d_n^{k+1}}{d_1^{k+1}}\right),$$

or $\lim_{k\to+\infty} \frac{d_i^{k+1}}{d_1^{k+1}} = 0, \forall i \neq 1$. On constate donc que dès que $\mathbf{x}_0$ n'est pas orthogonal au vecteur propre associé à la valeur propre de plus grand module, l'itération de A sur $\mathbf{x}_0$ va aligner le résultat sur la direction de ce vecteur propre. La généralisation à des matrices diagonalisables est assez simple. La généralisation à des matrices non nécessairement diagonalisables mais dont la valeur propre de plus grand module est unique est plus technique.

Pour finir, remarquons que la puissance itérée appliquée à $\mathbf{A}^{-1}$ permet de calculer la plus petite valeur propre (lorsqu'elle est la seule à réaliser le minimum en module). Lorsque l'unicité de la valeur propre réalisant le maximum en module, n'est plus vérifiée, cette méthode ne converge pas nécessairement. Des adaptations de la méthode sont possibles en cas

de symétrie de spectre ou plus généralement si des liens sur les valeurs propres des matrices sont connus (parfois identifiables dans la physique ou la biologie du problème modélisé).

2 LINEARISATION ET APPROXIMATION

2.1 Développement en série et développements limités

2.1.1 Fonctions de la variable réelle.

Nous nous bornerons à établir la formule de Taylor. Elle permet de développer localement une fonction sous la forme d'un polynôme et de borner l'erreur d'approximation.

Théorème des accroissements finis généralisés. Soient *f* et *g* deux fonctions réelles et continues sur un segment *[a,b]* et dérivables sur *]a,b[*, alors il existe un point $x \in]a,b[$ tel que

$$\big(f(b)-f(a)\big)g'(x)=\big(g(b)-g(a)\big)f'(x)\,.$$

Démonstration : Si on pose

$$h(t)=\big(f(b)-f(a)\big)g(t)-\big(g(b)-g(a)\big)f(t)$$

alors *h* est continue sur *[a,b]*, dérivable sur *]a,b[* et *h(a)=h(b)=f(b)g(a)-f(a)g(b)*. Pour établir le théorème il suffit de montrer que la fonction h' s'annule en un point $x \in]a,b[$. Si *h* est constante, c'est alors vrai pour tout point $x \in]a,b[$. Si $\exists t \in]a,b[, h(t) > h(a)$, alors soit *x* le point de *[a,b]* où *h* atteint son maximum, clairement $x \in]a,b[$ et $h'(x)=0$, ce qu'il fallait démontrer. Si $\exists t \in]a,b[, h(t) < h(a)$ nous raisonnons avec le minimum.

Pour obtenir le théorème des accroissements finis, nous posons simplement dans le théorème généralisé, *g(x)=x*.

Théorème des accroissements finis. Soit *f* une fonction réelle et continue sur un segment *[a,b]* et dérivable sur *]a,b[*, alors il existe un point $x \in]a,b[$ tel que : $f(b)-f(a)=(b-a)f'(x)$.

Théorème Formule de Taylor Soit *f* une fonction à valeurs réelles définie sur *[a,b]*, telle que *f* admette une dérivée d'ordre *n-1* continue sur *[a,b]* et telle que $f^{(n)}$ soit définie sur *]a,b[*, alors soit $x \in [a,b]$, on pose :

$$p(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!}(t-x)^k \ ,$$

soit $y \in [a,b]$ distinct de x alors il existe un point z, compris entre x et y, vérifiant

$$f(y) = p(y) + \frac{f^{(n)}(z)}{n!}(y-x)^n .$$

Nous admettons la formule de Taylor. Pour $n=1$, on reconnaît le théorème des accroissements finis. Dans le cas général, la formule de Taylor montre qu'une fonction f peut être approchée par un polynôme et que si $\left|f^{(n)}(t)\right|$ peut être majoré, alors l'erreur d'approximation peut être contrôlée. Par exemple, supposons que nous sachions que sur $[a,b]$ la dérivée seconde soit majorée $\left|f''(x)\right| < M$, alors au voisinage de x nous avons pour h de module suffisamment petit (pour que $x+h \in]a,b[$

$$f(x+h) = f(x) + f'(x)h + \frac{f''(z)h^2}{2}$$

soit

$$\left|f(x+h) - f(x) - f'(x)h\right| = \frac{\left|f''(z)\right|h^2}{2} \le \frac{Mh^2}{2} .$$

On voit donc que si M n'est pas trop grand, la tangente en x sera une bonne approximation de la fonction au voisinage de x (h petit).

2.1.2 Cas des fonctions de $\mathbb{R}^n$

Définition : soit $\mathbf{f}$ une application de $\mathbb{R}^n$ vers $\mathbb{R}^m$ alors on dit que $\mathbf{f}$ est différentiable en un point $\mathbf{x}$ de $\mathbb{R}^n$ s'il existe une application linéaire A de $\mathbb{R}^n$ vers $\mathbb{R}^m$ telle que

$$\lim_{\mathbf{h}\to 0} \frac{\left\|\mathbf{f(x+h)} - \mathbf{f(x)} - A\mathbf{h}\right\|}{\left\|\mathbf{h}\right\|} = 0$$

On dit alors que A est la différentielle de $\mathbf{f}$ en $\mathbf{x}$: $f'(\mathbf{x}) = A$

Exemple : différentielle d'une application linéaire. Soit A linéaire de $\mathbb{R}^n$ vers $\mathbb{R}^m$ alors $A(x+h) - Ax - Ah = 0$, donc $A'(\mathbf{x}) = A$.

Théorème : soit **f** une application de $\mathbb{R}^n$ vers $\mathbb{R}^m$ et **g** une application de $\mathbb{R}^m$ vers $\mathbb{R}^k$. Supposons que **f** soit différentiable en $\mathbf{x}_0$ et que **g** soit différentiable en $\mathbf{f}(\mathbf{x}_0)$ alors $\mathbf{h(x)} = \mathbf{g(f(x))}$ est différentiable en $\mathbf{x}_0$ et

$$\mathbf{h}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0)$$

La démonstration de ce théorème est laissée en exercice. Nous remarquons que la différentielle de la composition est le produit (c'est à dire la composition) des différentielles.

Considérons maintenant les bases canoniques $(\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n)$ de $\mathbb{R}^n$ et $(\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m)$ de $\mathbb{R}^m$. Les *m* fonctions coordonnées de **f** sont les fonctions réelles $(f_1, f_2, \cdots, f_m)$ définies par $\mathbf{f(x)} = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i$. La dérivée partielle de f_i par rapport à x_j est définie par :

$$D_j f_i(\mathbf{x}) = \lim_{t \to 0} \frac{f_i(\mathbf{x} + t\mathbf{e}_j) - f_i(\mathbf{x})}{t}$$

dès que cette limite existe. On note souvent $D_j f_i$ par $\frac{\partial f_i}{\partial x_j}$.

Théorème : Soit f une application d'un ouvert *O* de $\mathbb{R}^n$ vers $\mathbb{R}^m$ différentiable en un point $x \in O$. Alors, toutes les dérivées partielles $D_j f_i(\mathbf{x})$ existent et

$$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m} D_j f_i(\mathbf{x})\mathbf{u}_i .$$

La démonstration de ce théorème est laissée en exercice. La réciproque peut être fausse : l'existence des dérivées partielles n'entraîne pas nécessairement la différentiabilité.

Soit $\mathbf{M}_{\mathbf{f}'(\mathbf{x})}$ la matrice représentant l'application linéaire $\mathbf{f}'(\mathbf{x})$ dans les bases canonique de $\mathbb{R}^n$ et $\mathbb{R}^m$, alors $\mathbf{f}'(\mathbf{x})\mathbf{e}_j$ est la j^{ieme} colonne de $\mathbf{M}_{\mathbf{f}'(\mathbf{x})}$ et donc $\frac{\partial f_i}{\partial x_j}(\mathbf{x})$ est $\left(\mathbf{M}_{\mathbf{f}'(\mathbf{x})}\right)_{i,j}$. La matrice

$$\mathbf{M}_{\mathbf{f}'(\mathbf{x})} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

est appelée matrice Jacobienne de **f** en **x**. Si $\mathbf{h} = \sum h_j \mathbf{e}_j$, alors nous avons :

$$\mathbf{f}'(\mathbf{x})\mathbf{h} = \sum_{i=1}^{m}\left(\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(\mathbf{x})\right)h_j\right)\mathbf{u}_i \ .$$

Pour terminer, nous donnons, sans la démontrer, la formule de Taylor pour les fonctions de $\mathbb{R}^n$ vers $\mathbb{R}$, (nous pourrions l'appliquer aux fonctions coordonnées des fonctions de $\mathbb{R}^n$ vers $\mathbb{R}^m$) :

$$f(\mathbf{a}+\mathbf{x}) = \sum_{k=0}^{m-1}\frac{1}{k!}\sum (D_{i_1 i_2 \cdots i_k} f)(\mathbf{a}) x_{i_1} x_{i_2} \cdots x_{i_k} + r(\mathbf{x})$$

avec $\lim_{\mathbf{x}\to 0}\dfrac{r(\mathbf{x})}{\|\mathbf{x}\|^{m-1}} = 0$. La somme $\sum (D_{i_1 i_2 \cdots i_k} f)(\mathbf{a}) x_{i_1} x_{i_2} \cdots x_{i_k}$ est la somme suivant les n^k k-uplets d'entiers choisis parmi $1,2,\cdots,n$.

Par exemple, à l'ordre 2 : $f(\mathbf{a}+\mathbf{x}) = f(\mathbf{a}) + \sum_{j=1}^{n}(D_j f)(\mathbf{a}) x_j + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(D_{ij} f)(\mathbf{a}) x_i x_j + r(\mathbf{x})$. On note le vecteur $((D_j f)(\mathbf{a}))_{j=1,\cdots,n}$ le gradient de f en **a**, avec $(D_j f)(\mathbf{a}) = \dfrac{\partial f}{\partial x_j}(\mathbf{a})$, et la matrice **H** d'élément courant $H_{i,j} = (D_{ij} f)(\mathbf{a}) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$, le Hessien de f en **a**. Ainsi, on obtient :

$$f(\mathbf{a}+\mathbf{x}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{x}\rangle + \tfrac{1}{2}\langle \mathbf{x}, \mathbf{H}\mathbf{x}\rangle + r(\mathbf{x}) \ .$$

2.2 Développement en série de Fourier

Nous allons montrer que les fonctions $e_n(t) = e^{\frac{2i\pi nt}{a}}$, $n = -N,\cdots,N$, de la variable réelle t, avec $a>0$ et i tel que $i^2 = -1$, forment un système orthonormal de l'espace des polynômes trigonométriques de degré N, c'est à dire de l'ensemble des fonctions qui s'écrivent sous la forme : $p(t) = \sum_{n=-N}^{N} c_n e^{\frac{2i\pi nt}{a}}$, où $c_n \in \mathbb{C}$ est une suite de nombres complexes quelconques. Le produit scalaire est simplement défini par $f.g = \int_0^a f(t)\overline{g}(t)dt$

La question que l'on cherche à résoudre est la suivante : Toute fonction périodique *f(t)* de période *a,* admet-elle un développement en série de Fourier, c'est dire : $f(t)=\sum_{n=-\infty}^{\infty} c_n e^{\frac{2i\pi nt}{a}}$?

2.2.1 Orthogonalité des $e_n(t)$ et ses conséquences

Il est très facile de voir que

$$e_n.e_m=\int_0^a e_n(t)\overline{e}_m(t)dt=\begin{cases}0 & \text{si} \quad m\neq n\\ a & \text{sinon}\end{cases}$$

Notons $T_N=\left\{p, p(t)=\sum_{n=-N}^{N} c_n e^{\frac{2i\pi nt}{a}}\right\}$, l'ensemble des polynômes trigonométriques, alors nous pouvons déduire de l'orthogonalité des $e_n(t)$ que l'espace T_N est de dimension *2N+1* et que $\frac{1}{\sqrt{a}}e_n(t), n=-N,\cdots,N,$ est une base orthonormée de T_N. On obtient donc naturellement :

$p.e_m=\int_0^a p(t)\overline{e}_m(t)dt=\sum_{n=-N}^{N} c_n\int_0^a e_n\overline{e}_m(t)dt=ac_m$, donc on en déduit que

$$c_m=\frac{1}{a}\int_0^a p(t)\overline{e}_m(t)dt=\frac{1}{a}\int_0^a p(t)e^{\frac{-2i\pi nt}{a}}dt.$$

Cette formule, dite de Fourier, donne explicitement les coefficients de Fourier c_n en fonction de *p*. On obtient directement et de la même manière l'égalité de Parseval pour les polynômes trigonométriques : $\frac{1}{a}\|p\|^2=\frac{1}{a}\int_0^a |p(t)|^2dt=\sum_{n=-N}^{N}|c_n|^2$ (qui se lit dans l'autre sens !).

2.2.2 Développement en série de sinus et cosinus

On peut réécrire la série d'un polynôme trigonométrique en série de sinus et de cosinus sous la forme :

$$p(t)=\sum_{n=-N}^{N} c_n e^{\frac{2i\pi nt}{a}}=c_0+\sum_{n=1}^{N} c_n e^{\frac{2i\pi nt}{a}}+c_{-n}e^{\frac{-2i\pi nt}{a}}=c_0+\sum_{n=1}^{N} a_n\cos\left(\frac{2i\pi nt}{a}\right)+b_n\sin\left(\frac{2i\pi nt}{a}\right),$$

avec

$$\begin{cases}a_n &= & c_n+c_{-n}\\ b_n &= & i(c_n-c_{-n})\end{cases}.$$

En utilisant la formule de Fourier, il est aisé de constater que

$$a_n = \frac{2}{a}\int_0^a p(t)\cos\left(\frac{2i\pi nt}{a}\right)dt\,,\; b_n = \frac{2}{a}\int_0^a p(t)\sin\left(\frac{2i\pi nt}{a}\right)dt\,.$$

On constate aisément que si *p* est pair (*p(t)=p(-t)*) alors $\forall n, b_n = 0 \Leftrightarrow c_n = c_{-n}$; si *p* est impair (*p(t)=-p(-t)*) alors $\forall n, a_n = 0 \Leftrightarrow c_n = -c_{-n}$. En effet, comme *p(t)* et $e_n(t)$ sont périodiques nous avons $a_n = \frac{2}{a}\int_0^a p(t)\cos\left(\frac{2i\pi nt}{a}\right)dt = \frac{1}{a}\int_{-a/2}^{a/2} p(t)\cos\left(\frac{2i\pi nt}{a}\right)dt$ s'annule si *p(t)* est impair car puisque $\cos\left(\frac{2i\pi nt}{a}\right)$ est pair, le produit avec *p(t)* est impair et donc l'intégrale sur un domaine symétrique est nul.

2.2.3 Série de Fourier d'un signal périodique

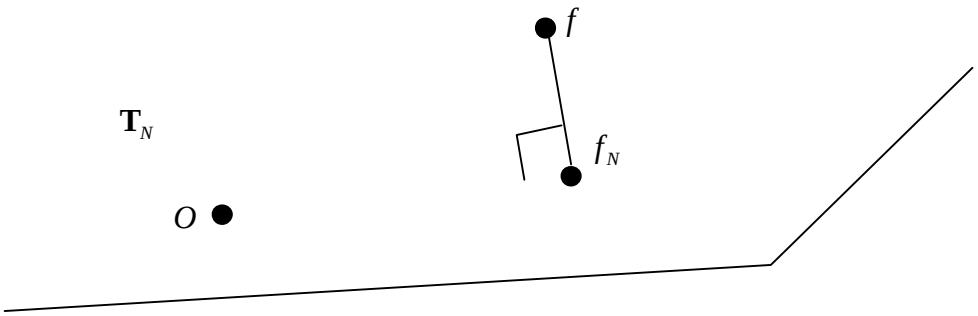
Nous souhaitons généraliser les résultats précédents à toute fonction périodique *f*. Nous aurons besoin que la fonction considérée *f(t)* soit de carré intégrable sur une période. Nous supposerons donc que $\|f\|^2 = \int_0^a |f(t)|^2 dt$ existe. Plus précisément, nous supposerons que $f \in L_P^2([0, a])$, l'espace de Hilbert des fonctions périodiques de période *a* de carré intégrable. Cet espace est un espace vectoriel, muni du produit scalaire (ou forme hermitienne dans le cas complexe) $f.g = \int_0^a f(t)\overline{g}(t)dt$, il est complet pour la norme qui en dérive $\|f\|_{L_P^2([0,a])} = \sqrt{f.f} = \sqrt{\int_0^a |f(t)|^2 dt}$. La construction de $L_P^2([0,a])$ a justement pour objectif d'obtenir un ensemble complet pour la norme précédente. En effet, dans un complet, on peut définir la notion de projection sur un fermé comme la limite d'une suite qui *converge dans* le fermé, en particulier la projection sur un sous-espace vectoriel de dimension fini.

En effet, nous souhaitons déterminer la meilleure approximation *dans* T_N de *f*. Il s'agit de trouver $f_N(t) = \sum_{n=-N}^{N} c_n e_n$, tel que :

$$\|f - f_N\|_{L_P^2([0,a])} = \min_{p \in T_N} \|f - p\|_{L_P^2([0,a])}\,.$$

La caractérisation de la projection sur un sous-espace vectoriel est

$$(f - f_N).\, p = 0, \forall p \in T_N\,.$$



Ainsi, avec $p = e_n$, nous avons $f.e_n = f_N.e_n, \forall n = -N, \cdots, N$, c'est à dire $c_n = c_n(f) = \frac{1}{a} f.e_n = \frac{1}{a}\int_0^a f(t)\overline{e}_n(t)dt$. On en déduit donc le théorème suivant :

Théorème. Il existe un polynôme trigonométrique f_N dans T_N, et un seul, tel que

$$\|f - f_N\|_{L_P^2([0,a])} = \min_{p \in T_N} \|f - p\|_{L_P^2([0,a])}$$

c'est $$f_N(t) = \sum_{n=-N}^{N} c_n e^{\frac{2i\pi nt}{a}}$$

avec $$c_n = c_n(f) = \frac{1}{a}\int_0^a f(t) e^{\frac{-2i\pi nt}{a}} dt .$$

Nous avons (Pythagore) :

$$\|f\|^2_{L_P^2([0,a])} = \|f_N\|^2_{L_P^2([0,a])} + \|f - f_N\|^2_{L_P^2([0,a])} = a\sum_{n=-N}^{N} |c_n|^2 + \|f - f_N\|^2_{L_P^2([0,a])},$$

qui nous permet d'obtenir l'inégalité de Bessel :

$$\sum_{n=-\infty}^{\infty} |c_n(f)|^2 \le \frac{1}{a}\int_0^a |f(t)|^2 dt .$$

Nous pouvons donc en déduire que la série $\sum_{n=-\infty}^{\infty}|c_n(f)|^2$ est convergente et donc $c_n(f)\to 0$, lorsque $|n|$ tend vers l'infini. Que devient f_N lorsque *N* tend vers l'infini ? On peut montrer (difficile et donc admis) que

Théorème. Si $f \in L_P^2([0,a])$, alors la meilleure approximation f_N de *f* dans $\mathbf{T}_N$ tend vers *f* dans $L_P^2([0,a])$ quand *N* tend vers l'infini, c'est à dire

$\|f - f_N\|^2_{L_P^2([0,a])} = \int_0^a |f(t) - f_N(t)|^2 dt \to 0$, quand $N \to +\infty$

L'égalité de Parseval

$$\sum_{n=-\infty}^{\infty}|c_n(f)|^2 = \tfrac{1}{a}\int_0^a |f(t)|^2 dt$$

est une conséquence directe du théorème précédent et de la relation de Pythagore. Elle est équivalente à :

$$\tfrac{1}{4}|a_0(f)|^2 + \tfrac{1}{2}\sum_{n=-\infty}^{\infty}|a_n(f)|^2 + |b_n(f)|^2 = \tfrac{1}{a}\int_0^a |f(t)|^2 dt\,.$$

L'unicité des coefficients de Fourier est une conséquence directe de l'égalité de Parseval. En effet, si $c_n(f) = c_n(g), \forall n$, alors par linéarité de l'intégrale, $c_n(f-g) = 0, \forall n$, et donc l'égalité de Parseval entraîne $\int_0^a |f(t) - g(t)|^2 dt$, soit *f*=*g* au sens de $L_P^2([0,a])$.

Enfin, on peut se demander ce qui se passe à *t* fixé ? Nous ne donnerons que deux résultats (admis) qui néanmoins contiennent l'idée générale suivante : *f* doit être suffisamment régulière pour que sa série de Fourier converge ponctuellement en t vers *f(t)* et plus *f* est régulière, plus rapide est la convergence.

Théorème de Dirichlet -Soit f une fonction périodique intégrable sur une période. Si elle admet en t_0 une limite à droite $f(t_0^+)$ et à gauche $f(t_0^-)$, de même qu'une dérivée à droite et à gauche en t_0 alors $f_N(t_0)$ tend vers $\tfrac{1}{2}\left(f(t_0^+) + f(t_0^-)\right)$, lorsque *N* tend vers l'infini.

On peut remarquer que si *f* est une fonction périodique de carré intégrable sur une période alors elle est intégrable.

Théorème. Si la fonction f est continue et dérivable et f' est continue, alors

- La série de Fourier de f' s'obtient en dérivant terme à terme celle de *f.*
- Les coefficients de Fourier de *f* vérifient $\sum_{n=-\infty}^{\infty} |c_n(f)| < +\infty$.
- La série de Fourier de *f* converge normalement (donc uniformément) sur $\mathbb{R}$ vers *f*.

La démonstration de ce théorème s'obtient par intégration par partie sur les coefficients de Fourier.

3 TRANSFORMATIONS

3.1 Transformée de Fourier

3.1.1 Cadre $L^1(\mathbb{R})$ et transformée de Fourier inverse

Nous voulons généraliser la notion de série de Fourier. En effet, les fonctions que nous manipulons ne sont pas nécessairement périodiques. Nous allons définir la transformation de Fourier d'une fonction *f*. Comme pour les coefficients de Fourier, il suffit a priori que la fonction soit intégrable $f \in L^1(\mathbb{R})$ pour que la définition suivante ait un sens.

$$\hat{f}(\xi) = \mathbf{F}f(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^a f(t) e^{-i\xi t} dt \text{ et } \tilde{f}(\xi) = \overline{\mathbf{F}} f(\xi) = \frac{1}{\sqrt{2\pi}} \int_0^a f(t) e^{i\xi t} dt$$

$\hat{f}(\xi)$ est la transformée de Fourier de *f*, $\tilde{f}(\xi)$ est la transformée de Fourier conjuguée de *f*. Nous allons voir dans la suite que l'inverse de la transformée de Fourier, lorsque cela a un sens, est la transformée de Fourier conjuguée : $\hat{\tilde{f}}(x) = \tilde{\hat{f}}(x) = f(x)$. Le problème essentiel est que de la transformée de Fourier d'une fonction n'est pas toujours intégrable et donc sa transformée de Fourier conjuguée n'existe pas en temps que fonction. Nous allons énoncer, sans les démontrer, les propriétés essentielles de la transformée de Fourier.

Théorème de Riemann-Lebegue Soit $f \in L^1(\mathbf{R})$, nous avons :

- $\hat{f}$ est une fonction continue et bornée sur $\mathbf{R}$.
- La transformée de Fourier est un opérateur linéaire et continu de $L^1(\mathbf{R})$ dans l'ensemble des fonctions continues bornées et $\left\|\hat{f}\right\|_\infty < \|f\|_1$.
- $\lim_{|\xi| \to +\infty} \left|\hat{f}(\xi)\right| = 0$.
- Soit $f \in L^1(\mathbf{R})$ et $g \in L^1(\mathbf{R})$ alors $\hat{f}g \in L^1(\mathbf{R})$ et $f\hat{g} \in L^1(\mathbf{R})$ et on a :

$$\int \hat{f}(t)g(t)dt = \int f(t)\hat{g}(t)dt$$

Nous avons bien entendu les mêmes propriétés pour la transformée de Fourier conjuguée.
Une propriété remarquable de la transformée de Fourier est de transformée la dérivation en la multiplication par un monôme.

1. Si $t^k f(t) \in L^1(\mathbf{R}), \forall k = 1,2,\cdots,n,$ alors $\hat{f}$ est n fois dérivable et on a :

$$\hat{f}^{(k)}(\xi) = \mathbf{F}\left((-ix)^k f(x)\right)(\xi), k = 1,2,\cdots,n,$$

2. Si $f \in L^1(\mathbf{R})$ est n fois dérivable et si toutes ses dérivées sont intégrables alors

$$\mathbf{F}\left(f^{(k)}\right)(\xi) = (i\xi)^k \hat{f}(\xi), k = 1,2,\cdots,n,$$

3. Si $f \in L^1(\mathbf{R})$ est à support borné alors $\hat{f}$ est infiniment dérivable.

Nous définissons l'opérateur de translation d'une fonction : $\mathbf{T}_a f(t) = f(t-a)$ .Nous pouvons vérifier en exercice la propriété remarquable que la transformée de Fourier de la translatée d'une fonction est la transformée de Fourier déphasée de a : $F\left(\mathbf{T}_a f\right)(\xi) = e^{-ia\xi}\hat{f}(\xi)$.

Théorème Si f et si $\hat{f}$ sont dans $L^1(\mathbf{R})$ alors on a $\hat{\hat{f}}(t) = \tilde{\hat{f}}(t) = f(t)$ en tout point où f est continue.

La démonstration technique est admise.

3.1.2 Transformée de Fourier dans $\mathcal{S}(\mathbf{R})$ et $L^2(\mathbf{R})$

Le cadre dans lequel les choses se "passent bien" pour la transformée de Fourier est celui des fonctions très régulières et à décroissance rapide. En effet, f est dite à décroissance rapide si $\forall p \in \mathbf{N}$, $\lim_{|x| \to \infty} \left|x^p f(x)\right| = 0$ et nous avons vu que la transformée de

Fourier d'une fonction à décroissance rapide était infiniment dérivable. De plus, si une fonction est infiniment dérivable et que ses dérivées sont intégrables alors sa transformée de Fourier est à décroissance rapide. On introduit $\mathcal{S}$ l'espace des fonctions infiniment dérivable, à décroissance rapide ainsi que toutes leurs dérivées. On peut montrer que la transformée de Fourier est une bijection isométrique (pour $\|f\| = \sqrt{\int f^2(t)dt}$) de $\mathcal{S}$ dans $\mathcal{S}$. L'inverse de la transformée de Fourier est bien entendu la transformée de Fourier inverse définie dans le paragraphe précédent. Ces résultats permettent de définir la transformée de Fourier dans $L^2(\mathbb{R})$ comme l'unique prolongement de la transformée de Fourier dans $\mathcal{S}$ ($\mathcal{S}$ est dense dans $L^2(\mathbb{R})$). Le résultat essentiel est alors que la transformée de Fourier est une bijection isométrique de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$.

REFERENCES

Gasquet C., Witomski P., *Analyse de Fourier et applications*, Masson (enseignement des Mathématiques),1995.

Lascaux P., Théodor R., *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, tome 1 et 2, Masson, 1986.

Rudin W., *Principes d'analyse Mathématique*. Ediscience international, Paris, 1995.