

CORRIGÉ

Agrégation interne · CORRIGÉ

CHAPITRES

DM n°06 — Projection orthogonale et méthode des moindres carrés

Orthogonalité et projection
orthogonale · Moindres carrés et
équations normales · Matrices de
Gram · Ajustement affine ·
Approximation de fonctions

I Ajuster trois points

1

Si les trois points appartenaient au graphe de $x \mapsto \alpha x + \beta$, on aurait

$$-\alpha + \beta = 1, \quad \beta = 1, \quad \alpha + \beta = 2.$$

La deuxième égalité donne $\beta = 1$ et la troisième $\alpha = 1$, ce qui contredit la première. Un ajustement exact est donc impossible.

2

On développe

$$\begin{aligned} S(\alpha, \beta) &= (-\alpha + \beta - 1)^2 + (\beta - 1)^2 + (\alpha + \beta - 2)^2 \\ &= 2\alpha^2 - 2\alpha + 3\beta^2 - 8\beta + 6. \end{aligned}$$

En complétant les carrés,

$$S(\alpha, \beta) = 2 \left(\alpha - \frac{1}{2} \right)^2 + 3 \left(\beta - \frac{4}{3} \right)^2 + \frac{1}{6}.$$

Ainsi

$$\alpha = \frac{1}{2}, \quad \beta = \frac{4}{3}$$

et

$$\min_{\mathbb{R}^2} S = \frac{1}{6}.$$

Le minimum est unique puisque les deux carrés doivent simultanément s'annuler.

3

(a) On a

$$\alpha X + \beta U = \begin{pmatrix} -\alpha + \beta \\ \beta \\ \alpha + \beta \end{pmatrix}.$$

Par conséquent

$$Y - (\alpha X + \beta U) = \begin{pmatrix} 1 + \alpha - \beta \\ 1 - \beta \\ 2 - \alpha - \beta \end{pmatrix},$$

et donc

$$\|Y - (\alpha X + \beta U)\|^2 = (-\alpha + \beta - 1)^2 + (\beta - 1)^2 + (\alpha + \beta - 2)^2 = S(\alpha, \beta).$$

Lorsque (α, β) parcourt $\mathbb{R}^2$, le vecteur $\alpha X + \beta U$ parcourt F . Minimiser S revient donc à chercher le point de F le plus proche de Y .

(b) Si $aX + bU = 0$, la deuxième coordonnée donne $b = 0$, puis la première donne $a = 0$. La famille (X, U) est donc libre, et

$$\boxed{\dim F = 2}.$$

(c) Avec $\alpha = \frac{1}{2}$ et $\beta = \frac{4}{3}$,

$$P = \frac{1}{2}X + \frac{4}{3}U = \begin{pmatrix} 5/6 \\ 4/3 \\ 11/6 \end{pmatrix},$$

puis

$$R = Y - P = \begin{pmatrix} 1/6 \\ -1/3 \\ 1/6 \end{pmatrix}.$$

On vérifie

$$\langle R, X \rangle = -\frac{1}{6} + \frac{1}{6} = 0, \quad \langle R, U \rangle = \frac{1}{6} - \frac{1}{3} + \frac{1}{6} = 0.$$

Comme X et U engendrent F , on obtient

$$\boxed{R \in F^\perp}.$$

4 _____

(a) Comme $Y - P = R \in F^\perp$ et $P - Z \in F$, les vecteurs $Y - P$ et $P - Z$ sont orthogonaux. Or

$$Y - Z = (Y - P) + (P - Z).$$

Le théorème de Pythagore donne donc

$$\boxed{\|Y - Z\|^2 = \|Y - P\|^2 + \|P - Z\|^2}.$$

(b) Pour tout $Z \in F$,

$$\|Y - Z\|^2 \geq \|Y - P\|^2,$$

avec égalité si et seulement si $P - Z = 0$, c'est-à-dire $Z = P$. Ainsi P est l'unique point de F le plus proche de Y . D'après la question 3(a),

$$\min S = \|Y - P\|^2 = \|R\|^2 = \frac{1}{36} + \frac{1}{9} + \frac{1}{36} = \boxed{\frac{1}{6}}.$$

II Projection orthogonale

5

Le vecteur nul appartient à $F^\perp$. Si $x, x' \in F^\perp$ et $\lambda, \mu \in \mathbb{R}$, alors, pour tout $y \in F$,

$$\langle \lambda x + \mu x', y \rangle = \lambda \langle x, y \rangle + \mu \langle x', y \rangle = 0.$$

Ainsi $\lambda x + \mu x' \in F^\perp$: $F^\perp$ est un sous-espace vectoriel de E .

Si $x \in F \cap F^\perp$, alors $x \in F$ et x est orthogonal à tout vecteur de F , en particulier à lui-même. Donc

$$\|x\|^2 = \langle x, x \rangle = 0,$$

d'où $x = 0$. Ainsi

$$F \cap F^\perp = \{0\}.$$

6

(a) Chaque e_i appartient à F , donc

$$q = \sum_{i=1}^r \langle x, e_i \rangle e_i \in F.$$

Pour $1 \leq j \leq r$,

$$\begin{aligned} \langle x - q, e_j \rangle &= \langle x, e_j \rangle - \sum_{i=1}^r \langle x, e_i \rangle \langle e_i, e_j \rangle \\ &= \langle x, e_j \rangle - \langle x, e_j \rangle = 0, \end{aligned}$$

car la famille $(e_1, \dots, e_r)$ est orthonormée. Comme elle engendre F , on a $x - q \in F^\perp$.

(b) La question précédente fournit une décomposition

$$x = q + (x - q), \quad q \in F, \quad x - q \in F^\perp.$$

Supposons

$$x = p + r' = p' + r'', \quad p, p' \in F, \quad r', r'' \in F^\perp.$$

Alors

$$p - p' = r'' - r'.$$

Le membre de gauche appartient à F , celui de droite à $F^\perp$; leur valeur commune appartient donc à $F \cap F^\perp = \{0\}$. Ainsi $p = p'$ et $r' = r''$. La décomposition est unique.

On obtient donc, pour tout $x \in E$, un unique $p \in F$ tel que $x - p \in F^\perp$. C'est le projeté orthogonal $P_F(x)$. En particulier,

$$E = F \oplus F^\perp.$$

7

(a) Soit $z \in F$. On écrit

$$x - z = (x - p) + (p - z).$$

Comme $x - p \in F^\perp$ et $p - z \in F$, les deux termes sont orthogonaux. Donc

$$\|x - z\|^2 = \|x - p\|^2 + \|p - z\|^2.$$

Il s'ensuit que $\|x - z\| \geq \|x - p\|$, avec égalité si et seulement si $z = p$. Ainsi p est l'unique élément de F minimisant la distance à x .

(b) Supposons $p \in F$ et

$$\|x - p\| \leq \|x - z\| \quad \text{pour tout } z \in F.$$

Soit $h \in F$. Pour tout $t \in \mathbb{R}$, $p + th \in F$, donc

$$\|x - p\|^2 \leq \|x - p - th\|^2.$$

En développant,

$$0 \leq -2t\langle x - p, h \rangle + t^2\|h\|^2 \quad \text{pour tout } t \in \mathbb{R}.$$

Si $h \neq 0$ et $\langle x - p, h \rangle \neq 0$, le choix

$$t = \frac{\langle x - p, h \rangle}{\|h\|^2}$$

rendrait le membre de droite strictement négatif. Donc $\langle x - p, h \rangle = 0$. Le cas $h = 0$ est immédiat. Ainsi $x - p$ est orthogonal à tout $h \in F$, donc

$$x - p \in F^\perp.$$

(c) Pour $p \in F$:

- $p = P_F(x)$ implique $x - p \in F^\perp$ par définition ;
- si $x - p \in F^\perp$, l'unicité de la question 6 donne $p = P_F(x)$;
- la question (a) montre que $p = P_F(x)$ implique que p est l'unique meilleur approximant ;
- si p minimise la distance à x , la question (b) donne $x - p \in F^\perp$, donc $p = P_F(x)$.

Les trois assertions sont donc équivalentes.

8

(a) D'après la formule de la question 6,

$$P_F(x) = \sum_{i=1}^r \langle x, e_i \rangle e_i,$$

la linéarité de P_F est immédiate. Si $x \in F$, alors $P_F(x) = x$; par conséquent

$$P_F^2 = P_F.$$

Tout $P_F(x)$ appartient à F , et tout élément de F est fixé par P_F , donc

$$\boxed{\operatorname{Im}(P_F) = F}.$$

Enfin,

$$P_F(x) = 0 \iff x = 0 + (x - 0), \quad x \in F^\perp \iff x \in F^\perp,$$

d'où

$$\boxed{\ker(P_F) = F^\perp}.$$

(b) Complétons $(e_1, \dots, e_r)$ en une base orthonormée $(e_1, \dots, e_r, e_{r+1}, \dots, e_n)$ de E . La projection fixe les r premiers vecteurs et annule les suivants. Sa matrice est donc

$$\boxed{\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}}.$$

(c) Écrivons

$$y = P_F(y) + y_\perp, \quad y_\perp \in F^\perp.$$

Comme $P_F(x) \in F$,

$$\langle P_F(x), y \rangle = \langle P_F(x), P_F(y) \rangle.$$

De même, en écrivant $x = P_F(x) + x_\perp$,

$$\langle x, P_F(y) \rangle = \langle P_F(x), P_F(y) \rangle.$$

Ainsi

$$\boxed{\langle P_F(x), y \rangle = \langle x, P_F(y) \rangle}.$$

Par ailleurs,

$$\|x\|^2 = \|P_F(x)\|^2 + \|x - P_F(x)\|^2,$$

d'où $\|P_F(x)\| \leq \|x\|$. Il y a égalité si et seulement si $x - P_F(x) = 0$, c'est-à-dire si et seulement si $x \in F$.

III Le problème des moindres carrés

9

(a) Lorsque x parcourt $\mathbb{R}^n$, le vecteur Ax parcourt exactement

$$\operatorname{Im} A = \{Ax ; x \in \mathbb{R}^n\}.$$

Ainsi

$$\Phi(x) = \|Ax - y\|^2 = \|y - Ax\|^2,$$

et minimiser Φ revient à chercher le point de $\operatorname{Im} A$ le plus proche de y .

(b) Comme $\text{Im } A$ est un sous-espace de $\mathbb{R}^m$, il existe un unique point

$$p = P_{\text{Im } A}(y)$$

minimisant la distance à y . Il existe au moins un x tel que $Ax = p$, puisque $p \in \text{Im } A$. Ainsi Φ admet un minimum. Si x et x' sont deux minimiseurs, alors

$$Ax = Ax' = p.$$

Le vecteur ajusté est donc unique, même si le paramètre ne l'est pas nécessairement.

(c) Soit x_0 un minimiseur. Un vecteur x est aussi minimiseur si et seulement s'il donne le même vecteur ajusté :

$$Ax = Ax_0 \iff A(x - x_0) = 0 \iff x - x_0 \in \ker A.$$

Par conséquent l'ensemble des minimiseurs est

$$\boxed{x_0 + \ker A}.$$

Il est réduit à un seul vecteur si et seulement si

$$\boxed{\ker A = \{0\}}.$$

10

(a) Pour $x = (x_j) \in \mathbb{R}^n$ et $z = (z_i) \in \mathbb{R}^m$,

$$\begin{aligned} \langle Ax, z \rangle &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) z_i \\ &= \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij} z_i \right) = \langle x, A^T z \rangle. \end{aligned}$$

(b) Pour $z \in \mathbb{R}^m$,

$$\begin{aligned} z \in (\text{Im } A)^\perp &\iff \langle Ax, z \rangle = 0 \quad \text{pour tout } x \in \mathbb{R}^n \\ &\iff \langle x, A^T z \rangle = 0 \quad \text{pour tout } x \in \mathbb{R}^n \\ &\iff A^T z = 0. \end{aligned}$$

Donc

$$\boxed{(\text{Im } A)^\perp = \ker(A^T)}.$$

(c) Le vecteur x minimise Φ si et seulement si Ax est le projeté orthogonal de y sur $\text{Im } A$, c'est-à-dire si et seulement si

$$y - Ax \in (\text{Im } A)^\perp.$$

D'après (b), cela équivaut à

$$A^T(y - Ax) = 0,$$

ou encore

$$A^T A x = A^T y.$$

Les équations normales traduisent donc exactement l'orthogonalité du résidu au sous-espace des vecteurs ajustés.

11

- (a) La (i, j) -ième entrée de $A^T A$ est le produit scalaire de la i -ième et de la j -ième colonne de A :

$$(A^T A)_{ij} = \langle a_i, a_j \rangle.$$

C'est donc la matrice de Gram de la famille $(a_1, \dots, a_n)$.

- (b) Si $x \in \ker A$, alors $A^T A x = 0$, donc $\ker A \subset \ker(A^T A)$. Réciproquement, si $A^T A x = 0$, alors

$$0 = \langle x, A^T A x \rangle = \langle A x, A x \rangle = \|A x\|^2,$$

donc $A x = 0$. Ainsi

$$\ker(A^T A) = \ker A.$$

- (c) $A^T A$ est une matrice carrée de taille n . Elle est inversible si et seulement si son noyau est réduit à $\{0\}$. D'après (b),

$$A^T A \text{ inversible} \iff \ker A = \{0\} \iff (a_1, \dots, a_n) \text{ est libre.}$$

- (d) Sous ces hypothèses, les équations normales ont l'unique solution

$$\hat{x} = (A^T A)^{-1} A^T y.$$

Le vecteur ajusté est donc

$$A \hat{x} = A(A^T A)^{-1} A^T y = P_{\text{Im } A}(y).$$

La matrice de la projection orthogonale sur $\text{Im } A$ est ainsi

$$A(A^T A)^{-1} A^T.$$

12

- (a) Si $x = (x_1, x_2)^T$, alors

$$A x = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix}$$

et

$$\|A x - y\|^2 = (x_1 + x_2 - 1)^2 + 1.$$

Le minimum vaut 1 et il est atteint exactement lorsque

$$x_1 + x_2 = 1.$$

Les minimiseurs sont donc les vecteurs

$$x = \begin{pmatrix} t \\ 1-t \end{pmatrix}, \quad t \in \mathbb{R}.$$

(b) Pour tout minimiseur,

$$Ax = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad y - Ax = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Le vecteur ajusté est unique, contrairement au paramètre x .

(c) On calcule

$$A^T A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad A^T y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Les équations normales sont donc

$$\begin{cases} x_1 + x_2 = 1, \\ x_1 + x_2 = 1. \end{cases}$$

Elles caractérisent exactement la droite affine de minimiseurs trouvée en (a). La singularité de $A^T A$ n'empêche donc pas les équations normales de caractériser les minimiseurs ; elle empêche seulement ici leur unicité.

IV Ajuster une droite à N observations

13

(a) Il suffit de prendre

$$A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix}.$$

Ses colonnes sont X et $\mathbf{1}$, et

$$A \begin{pmatrix} a \\ b \end{pmatrix} = aX + b\mathbf{1}.$$

(b) Les colonnes sont liées si et seulement s'il existe $(\lambda, \mu) \neq (0, 0)$ tel que

$$\lambda X + \mu \mathbf{1} = 0.$$

Si $\lambda = 0$, alors $\mu = 0$, contradiction. Il faut donc $\lambda \neq 0$, et

$$x_i = -\frac{\mu}{\lambda}$$

pour tout i : les x_i sont tous égaux. Réciproquement, s'ils sont tous égaux à c , alors $X = c\mathbf{1}$, donc les colonnes sont liées. Ainsi elles sont indépendantes si et seulement si les x_i ne sont pas tous égaux.

(c) Par hypothèse, les colonnes de A sont indépendantes, donc $\ker A = \{0\}$. D'après la partie III, le problème de moindres carrés possède alors un unique paramètre minimisant. Il existe donc un unique couple (a, b) recherché.

14

(a) Les équations normales s'écrivent

$$\begin{cases} \sum_{i=1}^N x_i(ax_i + b - y_i) = 0, \\ \sum_{i=1}^N (ax_i + b - y_i) = 0. \end{cases}$$

De façon équivalente,

$$\begin{cases} \sum_{i=1}^N x_i(y_i - ax_i - b) = 0, \\ \sum_{i=1}^N (y_i - ax_i - b) = 0. \end{cases}$$

(b) La deuxième équation donne

$$N\bar{y} - aN\bar{x} - Nb = 0,$$

donc

$$\boxed{b = \bar{y} - a\bar{x}}.$$

(c) En remplaçant b dans la première équation sous la forme utilisant le résidu,

$$\sum_{i=1}^N x_i[(y_i - \bar{y}) - a(x_i - \bar{x})] = 0.$$

Or

$$\sum_{i=1}^N (y_i - \bar{y}) = 0, \quad \sum_{i=1}^N (x_i - \bar{x}) = 0.$$

On peut donc remplacer x_i par $x_i - \bar{x}$ dans la somme, ce qui donne

$$\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) - a \sum_{i=1}^N (x_i - \bar{x})^2 = 0.$$

Comme les x_i ne sont pas tous égaux,

$$\sum_{i=1}^N (x_i - \bar{x})^2 > 0.$$

Ainsi

$$\boxed{a = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}}, \quad \boxed{b = \bar{y} - a\bar{x}}.$$

(d) Pour $x = \bar{x}$, la valeur de la droite ajustée est

$$a\bar{x} + b = a\bar{x} + \bar{y} - a\bar{x} = \bar{y}.$$

Elle passe donc par

$$\boxed{(\bar{x}, \bar{y})}.$$

15 _____

(a) On a

$$\langle X_c, \mathbf{1} \rangle = \sum_{i=1}^N (x_i - \bar{x}) = \sum_{i=1}^N x_i - N\bar{x} = 0.$$

Donc $X_c \perp \mathbf{1}$. Par ailleurs,

$$X_c = X - \bar{x} \mathbf{1}, \quad X = X_c + \bar{x} \mathbf{1},$$

d'où

$$\boxed{\text{Vect}(X, \mathbf{1}) = \text{Vect}(X_c, \mathbf{1})}.$$

(b) Les vecteurs X_c et $\mathbf{1}$ sont orthogonaux et non nuls. La projection de Y sur leur plan engendré est donc

$$P(Y) = \frac{\langle Y, X_c \rangle}{\|X_c\|^2} X_c + \frac{\langle Y, \mathbf{1} \rangle}{\|\mathbf{1}\|^2} \mathbf{1}.$$

Or

$$\frac{\langle Y, \mathbf{1} \rangle}{\|\mathbf{1}\|^2} = \frac{\sum_i y_i}{N} = \bar{y},$$

et

$$\frac{\langle Y, X_c \rangle}{\|X_c\|^2} = \frac{\sum_i (x_i - \bar{x}) y_i}{\sum_i (x_i - \bar{x})^2} = \frac{\sum_i (x_i - \bar{x}) (y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} = a.$$

Ainsi

$$P(Y) = aX_c + \bar{y} \mathbf{1} = aX + (\bar{y} - a\bar{x}) \mathbf{1},$$

ce qui redonne $b = \bar{y} - a\bar{x}$ et la formule de a de la question 14.

(c) Pour les trois points de la partie I,

$$\bar{x} = 0, \quad \bar{y} = \frac{1+1+2}{3} = \frac{4}{3}, \quad X_c = X.$$

Alors

$$\alpha = \frac{\langle Y, X \rangle}{\|X\|^2} = \frac{-1+2}{2} = \frac{1}{2},$$

et

$$\beta = \bar{y} - \alpha \bar{x} = \frac{4}{3}.$$

On retrouve bien le résultat de la question 2.

Enfin $X \perp U$ puisque

$$\langle X, U \rangle = -1 + 0 + 1 = 0.$$

Dans

$$\|Y - \alpha X - \beta U\|^2,$$

le terme croisé en $\alpha\beta$ est proportionnel à $\langle X, U \rangle$; il est donc nul.

V Approximer des fonctions

16

(a) La bilinéarité et la symétrie résultent de celles de l'intégrale et du produit réel. De plus,

$$\langle f, f \rangle = \int_{-1}^1 f(t)^2 dt \geq 0.$$

Si cette intégrale est nulle, la fonction continue f^2 , positive ou nulle, a une intégrale nulle. Elle est donc identiquement nulle : sinon elle serait strictement positive sur un petit intervalle par continuité, ce qui donnerait une intégrale strictement positive. Ainsi $f = 0$. La forme donnée est donc bien un produit scalaire.

(b) Les fonctions $t \mapsto t$ et $t \mapsto 1$ sont linéairement indépendantes et engendrent F , donc

$$\boxed{\dim F = 2}.$$

La construction de la question 6 n'utilise que l'existence d'une base orthonormée du sous-espace sur lequel on projette. Comme F est de dimension finie, on peut choisir une telle base dans F et, pour tout $f \in H$, définir

$$P_F(f) = \sum_i \langle f, e_i \rangle e_i.$$

Le même calcul que dans la question 6 montre alors que $f - P_F(f) \in F^\perp$. La dimension finie de l'espace ambiant n'est donc pas nécessaire ici.

(c) Écrivons $p(t) = at + b$. La condition de projection impose

$$t^2 - at - b \perp 1 \quad \text{et} \quad t^2 - at - b \perp t.$$

La première condition donne

$$\int_{-1}^1 (t^2 - at - b) dt = \frac{2}{3} - 2b = 0,$$

donc $b = \frac{1}{3}$. La seconde donne

$$\int_{-1}^1 t(t^2 - at - b) dt = -a \int_{-1}^1 t^2 dt = -\frac{2a}{3} = 0,$$

donc $a = 0$. Ainsi

$$p(t) = \frac{1}{3}.$$

La valeur minimale est

$$\begin{aligned} \int_{-1}^1 \left(t^2 - \frac{1}{3}\right)^2 dt &= \int_{-1}^1 \left(t^4 - \frac{2}{3}t^2 + \frac{1}{9}\right) dt \\ &= \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \boxed{\frac{8}{45}}. \end{aligned}$$

17

(a) La condition $f - p \in F^\perp$ signifie que, pour chaque $i \in \{0, \dots, d\}$,

$$0 = \langle f - p, \varphi_i \rangle = \langle f, \varphi_i \rangle - \sum_{j=0}^d c_j \langle \varphi_j, \varphi_i \rangle.$$

Comme le produit scalaire est symétrique,

$$\sum_{j=0}^d G_{ij} c_j = D_i,$$

avec

$$G_{ij} = \langle \varphi_i, \varphi_j \rangle, \quad D_i = \langle f, \varphi_i \rangle.$$

On obtient bien

$$\boxed{Gc = D}.$$

(b) Pour $c = (c_0, \dots, c_d)^T$,

$$\begin{aligned} c^T Gc &= \sum_{i,j=0}^d c_i c_j \langle \varphi_i, \varphi_j \rangle \\ &= \left\langle \sum_{i=0}^d c_i \varphi_i, \sum_{j=0}^d c_j \varphi_j \right\rangle \\ &= \left\| \sum_{j=0}^d c_j \varphi_j \right\|^2. \end{aligned}$$

Si $c \neq 0$, l'indépendance linéaire de $(\varphi_0, \dots, \varphi_d)$ implique

$$\sum_{j=0}^d c_j \varphi_j \neq 0,$$

donc $c^T Gc > 0$. La matrice symétrique G est définie positive ; en particulier son noyau est réduit à $\{0\}$ et elle est inversible.

18

(a) Si la famille est orthonormée, alors $G = I_{d+1}$. Le système de la question 17 devient

$$c_j = \langle f, \varphi_j \rangle, \quad 0 \leq j \leq d.$$

(b) Prenons pour φ_j les colonnes a_j de A . Alors

$$G_{ij} = \langle a_i, a_j \rangle,$$

donc $G = A^T A$. Pour $f = y$,

$$D_i = \langle y, a_i \rangle,$$

donc $D = A^T y$. Le système $Gc = D$ devient

$$A^T A c = A^T y,$$

qui est précisément le système des équations normales.

(c) Le sous-espace

$$F = \mathbb{R}_d[X]$$

est de dimension finie $d + 1$. La construction de la projection orthogonale sur un sous-espace de dimension finie s'applique donc dans $C([-1, 1], \mathbb{R})$ muni du produit scalaire de la question 16. Pour toute fonction continue f , il existe ainsi un unique

$$p = P_F(f) \in \mathbb{R}_d[X]$$

qui minimise $\|f - p\|$.

Dans la base monomiale $(1, X, \dots, X^d)$, les coefficients doivent être obtenus en résolvant le système de Gram $Gc = D$. Dans une base orthonormée $(\varphi_0, \dots, \varphi_d)$, ce système devient l'identité et

$$p = \sum_{j=0}^d \langle f, \varphi_j \rangle \varphi_j.$$

La même géométrie gouverne donc l'ajustement discret et l'approximation de fonctions : le meilleur approximant est caractérisé par l'orthogonalité du résidu au sous-espace d'approximation.