MAT 3741 : Algèbre linéaire appliquée, Hiver 2017
Devoir #3

À remettre ~~le vendredi le 3 mars~~ mardi le 7 mars à 10h.

Cette fois, tout le devoir est ensemble dans une partie mais vous êtes priés d'utiliser MATLAB pour faire les calculs (questions notés *), et pouvez remettre les résultats en-ligne au lieu de les recopier sur papier (espéciallement si Matlab ne vous donne que des résultats arrondis).

1. Soit A une matrice de taille $n \times m$ telle que $\mathrm{rang}(A) = m$.
   (a) Démontrer que $\overline{A^T}A$ est une matrice inversible, en utilisant la décomposition Q-R de A.
   (b) Démontrer que la projection orthogonale de $\mathbb{R}^n$ sur l'espace engendré par les colonnes de A est donnée par
   $$P = A(\overline{A^T}A)^{-1}\overline{A^T}.$$
   Il suit qu'on peut calculer la projection orthgonale sans faire Gram-Schmidt !
   (c) * Soit $A = \begin{bmatrix} 1 & i & 1 \\ 0 & i & -1 \\ i & 1 & -i \\ 0 & i & 1 \end{bmatrix}$. Calculer $(\overline{A^T}A)^{-1}$ et puis calculer P de deux façons : avec la formule en (b), et en calculant $A = QR$ une décomposition Q-R normalisée, or $P = Q\overline{Q^T}$. Quelques remarques de Matlab :

```
A'   est la conjuguée complexe de la transposée de A
inv(A) est l'inverse de A
[Q,R]=qr(A,0) donne notre décomposition QR, si rang(A)=#col(A)
norm(Pn-P,1)  donne la norme de la matrice Pn-P.  Quand il s'agit
              d'erreurs d'arrondissements, c'est mieux de calculer
              la norme de la différence que d'espérer que la différence
              sera la matrice nulle
```

2. Soit A une matrice de taille $n \times m$ de rang k quelconque, telle que $A = QR$ est sa décomposition Q-R normalisée. La pseudo-inverse de Moore-Penrose est la matrice
   $$A^+ = \overline{R^T}(R\overline{R^T})^{-1}\overline{Q^T}.$$
   (a) Démontrer, avec l'aide de (1), que cette formule est bien définie, et donner la taille de A^+.
   (b) Démontrer que A^+A est la projection orthgonale sur $\mathrm{Col}(\overline{A^T})$.
   (c) Démontrer que AA^+ est la projection orthogonale sur $\mathrm{Col}(A)$.
   (d) Dans quel circonstances est-ce que A^+ est une inverse à gauche, ou à droite, de A ?
   (e) * Soit A comme en (1c). Déterminer A^+ et comparer avec le résultat de la fonction pinv(A) de Matlab (en norme). [1]

3. Soit A une matrice de taille $n \times m$ de rang k quelconque.
   (a) Soit $y \in \mathrm{Col}(A)$ et posons $S = \{x \in \mathbb{R}^m \mid Ax = y\}$. Démontrer que l'élément de S ayant la plus petite norme est la projection de S sur $\mathrm{Col}(\overline{A^T})$. Indice : décrire S en termes de $\mathrm{Noy}(A)$ et puis identifier $\mathrm{Noy}(A)^{\perp}$. (Indice : Si les conjuguées complexes vous mêlent, oubliez-les en premier lieu.)

1. Par contre, avec MATLAB [Q0,R0]=qr(A) donne une matrice Q0 carrée, et de plus, il le fait d'une manière plus stable que ce qu'on fait à la main. Si vous choisissez une matrice A avec $\mathrm{rang}(A) = \dim(Col(A))$ alors vous pouvez récupérer notre matrice Q des premières $\mathrm{rang}(A)$ colonnes de $Q0$; mais si A n'est **pas** de plein rang, vous ne pouvez pas en générale récuperer notre décomposition Q-R normalisée. Points bonis si vous trouvez, ou produisez, une fonction Matlab qui retournera notre décomposition Q-R dans le cas générale de $n \times m$ de rang k.

(b) Démontrer (avec les résultats ci-haut) que pour tout $b \in \mathbb{R}^n$, si on pose $x = A^+b$ alors c'est un choix idéale : (i) la norme de $Ax - b$ est le minimum possible,[2] et (ii) la norme de x est le minimum possible tel que (i) soit vrai.

4. La regression statistique peut être défini comme suit. Soit

$$S = \{(x_{j,1}, x_{j,2}, \cdots, x_{j,n}, y_j) = (\vec{x_j}, y_j) \in \mathbb{R}^{n+1} \mid 1 \le j \le m\}$$

un ensemble de données. On veut trouver des scalaires $a_i \in \mathbb{R}$ tels que la fonction

$$g(x_1, \cdots, x_n) = \sum_{i=1}^{n} a_i x_i$$

ait la propriété que si on pose $y = (y_1, \cdots, y_m)$ et $z = (g(\vec{x_1}), \cdots, g(\vec{x_m}))$ alors

$$\|y - z\|$$

est minimale[3] parmis tous les choix possibles de scalaires a_i.

(a) Un cas particulier, pour fixer ses idées : Vos données sont

x_1	x_2	y
.5	.6	.5
.3	.4	.4
.9	.9	.8
.6	.6	.7

Vous voulez trouver a et b tel que $ax_1 + bx_2 = y$, pour chacune des quatres triplets (x_1, x_2, y) donnée dans ce tableau. Veuillez écrire ce système linéaire. Veuillez vous assurer que ce système est incohérent. Donc on veut une solution moindre carrées ; trouvez-la.[4] On appelle les valeurs $ax_1 + bx_2$ (pour (x_1, x_2) dans notre tableau) nos *prédictions* ; elles ne seront pas égale à nos *observations* qui sont les y. Leur différences sont les résiduels. Déterminer la norme du vecteur (taille 4×1) de résiduels dans ce cas. C'est une mesure de l'efficacité de notre prédiction ; en (d) on mentionne une meilleure.
Donc en générale (deux questions très rapides (b) et (c)) :

(b) Quelle est la matrice A telle que si on pose $u = (a_1, \cdots, a_n)$ alors

$$z = Au?$$

(c) Et : quel est le système linéaire dont la solution moindre carrrées est le vecteur de constantes u qu'on cherche ?

(d) * On vous donne un fichier excel avec des données. Vous pouvez facilement copier-coller ce que vous voulez dans MATLAB pour créer une matrice, genre

```
>> A= [ (paste data
from excel)
 ]
>> y =[ (paste data
from excel)
 ]
```

Trouvez la fonction linéaire $g(x_1, x_2, x_3) = ax_1 + ax_2 + ax_3 + b$ pour laquelle $\|\vec{y} - g(\vec{x})\|$ est minimale, en trouvant la solutions moindres carrées du système comme en (a) ci-haut. Indice : ajouter une colonne de 1 comme un x_4 pour obtenir la possibilité d'une constante.

---

2. Donc, en particulier c'est une solution moindre carrées si le système est incohérent, et une solution exacte autrement.

3. par rapport à la 2-norme

4. On a vu au cours que si $Bx = b$ est incohérent, alors $\overline{B^T}Bx = \overline{B^T}b$ est cohérent, et sa solution x_0 est dite la solution moindre carrées du système originale, puisque $\|Bx_0 - b\|_2$ est le minimum possible.

Puis calculer le *coéfficient de détermination* R^2 : Soit $\overline{y}$ la moyenne des valeurs $y_1, \cdots, y_n$ et poser $\vec{Y} = (\overline{y}, \overline{y}, \cdots, \overline{y})$ (n composantes). Alors R^2 est lié au rapport entre la norme de la différence entre les observations et les prédictions, et la variance totale S_{tot} des observations :[5]

$$R^2 = 1 - \frac{S_{res}}{S_{tot}} = 1 - \frac{\|g(\vec{x}) - \vec{y}\|^2}{\|\vec{y} - \vec{Y}\|^2}.$$

Typiquement $0 \leq R^2 \leq 1$; le plus proche que le résultat est à 1, le plus que les données sont bien modélisés par la fonction g.[6] Qu'en pensez-vous de ces données ?

Quelques commandes utiles :

```
A(:,4) = 1                 :  cré une quatrième colonne ayant que des 1
ybar = mean(y)             :  la moyenne
Stot = norm(y(:)-ybar)^2   :  y(:)-ybar soustrait ybar de chaque composante de y
```

5. The coefficient of determination of a linear regression is related to the quotient of the norm-squared of the residues (difference of data and prediction) and the variance of the observations, and is one measure of the goodness of fit of a regression.

6. Si vous étudiez en statistique, vous auriez des formules plus sophistiquées à appuyer ici ; veuillez comparer !