Décomposition QR. Méthode de Householder

L'espace vectoriel $\mathbb{R}^n$ est muni de sa structure canonique d'espace euclidien, pour laquelle la base canonique $(e) = e_1, e_2, \ldots, e_n$ est orthonormée.

On notera $[x]$ la matrice colonne des coordonnées d'un vecteur x de $\mathbb{R}^n$.

On rappelle que pour tous vecteurs x et y, $< x, y >= {}^{\mathrm{T}}[x][y]$.

Partie I

Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. On appelle *décomposition QR* de A tout couple (Q, R) de matrices de $\mathcal{M}_n(\mathbb{R})$ telles que $A = QR$, Q étant une matrice orthogonale (un élément du groupe $O(n)$) et R une matrice triangulaire supérieure.

Dans toute la suite, on suppose que A est inversible (il en est donc de même de R).

1. En invoquant le procédé d'orthogonalisation de Schmidt, montrer l'existence de la décomposition $A = QR$ si on suppose que les coefficients diagonaux de la matrice R sont strictement positifs. [S]
2. Par un raisonnement purement matriciel, montrer alors l'existence et l'unicité de la décomposition QR de A quand on fixe à l'avance et de manière quelconque le signe de chacun des coefficients diagonaux de R. [S]

Partie II

Dans cette partie, on étudie une métode permettant de trouver une décomposition QR d'une matrice inversible A, et utilisant des produits successifs par des matrices de réflexions (dites *matrices de Householder*).

1. (a) Comment s'écrit la réflexion (c'est-à-dire la symétrie orthogonale) h par rapport à l'hyperplan orthogonal à un vecteur non nul a ? [S]

   (b) Montrer que sa matrice dans la base canonique est $H = I_n - \dfrac{2}{\|a\|^2}[a]\,{}^{\mathrm{T}}[a]$.
   On dit que H est la *matrice de Householder* associée à a. [S]

   (c) Montrer que si u et v sont deux vecteurs distincts de $\mathbb{R}^n$ et de même norme, alors il existe une réflexion unique h telle que $h(u) = v$. [S]

   (d) Etudier l'existence d'une réflexion h transformant un vecteur u de $\mathbb{R}^n$ en $v = \lambda e_1$ colinéaire à e_1. On illustrera la situation par un schéma. [S]

   (e) Plus généralement, soit i un indice de $\{1, \ldots, n\}$, et $u = (x_1, x_2, \ldots, x_n)$.

   On pose $v = (x_1, \ldots, x_{j-1}, \lambda, 0, \ldots, 0)$ et $w = (0, \ldots, 0, x_j, \ldots, x_n)$.

   Etudier l'existence d'une réflexion h de $\mathbb{R}^n$ transformant u en v.

   Que dire des vecteurs $e_1, \ldots, e_{j-1}$ pour l'application h ? [S]

2. Dans la suite de cette question, et pour des commodités de notation, on ajoutera la matrice identité I_n à l'ensemble des matrices de Householder.

   (a) Déduire de ce qui précède qu'il existe $n-1$ matrices de Householder $H_1, H_2, \ldots, H_{n-1}$ telles que pour tout j les coefficients sous-diagonaux des j premières colonnes de $A_j = H_j H_{j-1} \cdots H_1 A$ soient nuls. [S]

   (b) Montrer comment des produits successifs par les matrices H_j permettent d'obtenir séparément les composantes Q et R d'une décomposition QR de la matrice A. [S]

   (c) Illustrer cette méthode en décomposant $A = \begin{pmatrix} -4 & 20 & 35 & 5 \\ -4 & -30 & -15 & 55 \\ -8 & 40 & -80 & -65 \\ 23 & -15 & 30 & 15 \end{pmatrix}$ [S]

Corrigé du problème

Partie I

1. La matrice A étant inversible, ses vecteurs colonnes $a_1, a_2, \ldots, a_n$ (considérés comme éléments de $\mathbb{R}^n$) forment une base (a) de $\mathbb{R}^n$, notée (a).

   A est alors la matrice de passage $P_{e,a}$ de la base canonique (e) à la base (a).

   Soient Q et R deux matrices de $\mathcal{M}_n(\mathbb{R})$, telles que $A = QR$.

   Ces deux matrices sont nécessairement inversibles ($\det Q \det R = \det A \neq 0$).

   Les vecteurs colonnes $q_1, q_2, \ldots, q_n$ de Q forment donc une base (q) de $\mathbb{R}^n$, et Q est la matrice de passage $P_{e,q}$ de la base (e) à la base (q).

   L'égalité $A = QR$ équivaut alors à $R = Q^{-1}A = P_{q,e}P_{e,a} = P_{q,a}$.

   Ainsi R est la matrice de passage de la base (q) à la base (a).

   Dans ces conditions :

   – Dire que Q est orthogonale, c'est dire que (q) est une base orthonormée de $\mathbb{R}^n$.

   – Dire que R est triangulaire supérieure à coefficients diagonaux positifs, c'est dire que, pour tout entier k compris entre 1 et n, a_k est une combinaison linéaire de $q_1, q_2, \ldots, q_k$, sa composante sur q_k étant strictement positive.

   On reconnait là les conditions décrivant l'orthonormalisation de la famille $a_1, a_2, \ldots, a_n$ par le procédé de Schmidt.

   On sait que ce procédé est constructif. Il prouve ainsi l'existence d'une famille $q_1, q_2, \ldots, q_n$ (donc une matrice Q et une matrice R), ce qui établit le résultat. [Q]

2. – Existence :

   Soit $A = Q^+R^+$ la décomposition évoquée dans la question précédente (les coefficients diagonaux de R étant strictement positifs.)

   Soit X un sous-ensemble quelconque de $\{1, \ldots, n\}$.

   Soit J la matrice diagonale dont les coefficients diagonaux successifs λ_i valent 1 si i est dans X et -1 sinon : c'est une matrice orthogonale qui vérifie $J^2 = I_n$.

   L'égalité $A = Q^+R^+$ s'écrit donc aussi $A = QR$ avec $Q = Q^+J$ et $R = JR^+$.

   La matrice Q est orthogonale (elle se déduit de Q^+ en inversant les signes dans les colonnes d'indice $j \notin X$) et R est triangulaire supérieure (elle se déduit de R^+ en inversant les signes dans les lignes d'indice $i \notin X$).

   On a ainsi formé une décomposition $A = QR$ de A pour laquelle le signe des coefficients diagonaux de R est fixé à l'avance.

   Il reste à prouver l'unicité d'une telle décomposition.

   – Unicité :

   Supposons que la matrice A possède deux décompositions, $A = Q_1R_1$ et $A = Q_2R_2$, les coefficients diagonaux respectifs de R_1 et de R_2 ayant le même signe.

   Il faut montrer $Q_1 = Q_2$ et $R_1 = R_2$.

   L'égalité $Q_1R_1 = Q_2R_2$ donne $Q_2^{-1}Q_1 = R_2R_1^{-1}$.

La matrice $M = R_2 R_1^{-1}$ est, tout comme R_1 et R_2, triangulaire supérieure.

Ses coefficients diagonaux sont les quotients deux à deux de ceux de R_2 et R_1 et sont donc strictement positifs.

Puisque $O(n)$ est un groupe, la matrice $M = Q_2^{-1} Q_1$ est une matrice orthogonale.

La matrice M^{-1} est donc à la fois triangulaire supérieure et triangulaire inférieure (comme inverse et comme transposée d'une matrice ayant cette propriété.) Les matrices M et M^{-1} sont donc diagonales.

Du fait de l'égalité $M^{-1} = {}^{\mathrm{T}}M$, chaque coefficient diagonal de M est égal à son propre inverse et vaut donc à priori ± 1.

Mais on sait que ces coefficients sont positifs : ils sont donc égaux à 1.

On en déduit $M = I_n$, c'est-à-dire $Q_1 = Q_2$ et $R_1 = R_2$.

On a ainsi prouvé l'existence et l'unicité de la décomposition $A = QR$ (sous réserve d'existence) dans le cas où on impose à l'avance le signe de chacun des coefficients diagonaux de R.

Au total, on peut donc considérer que A possède 2^n décompositions $A = QR$, chacune d'elles se déduisant de la décomposition $A = Q^+ R^+$ en inversant le signe d'une ou de plusieurs colonnes de Q^+ et celui des lignes de même indice dans R^+.

[Q]

Partie II

1. (a) Cette réflexion est définie par : $\forall u \in \mathbb{R}^n, h(u) = u - \dfrac{2}{\|a\|^2} < a, u > a$. [Q]

   (b) Pour tout vecteur u de $\mathbb{R}^n$,

   $$[h(u)] = [u] - \frac{2}{\|a\|^2}\,{}^{\mathrm{T}}[a][u]\,[a] = [u] - \frac{2}{\|a\|^2}\,[a]\,{}^{\mathrm{T}}[a]\,[u] = \Big(I_n - \frac{2}{\|a\|^2}\,[a]\,{}^{\mathrm{T}}[a]\Big)[u]$$

   Remarque : on a pu écrire ${}^{\mathrm{T}}[a][u]\,[a] = [a]\,{}^{\mathrm{T}}[a]\,[u]$ car ${}^{\mathrm{T}}[a][u]$ est un scalaire.

   Le calcul précédent montre bien que $H = I_n - \dfrac{2}{\|a\|^2}\,[a]\,{}^{\mathrm{T}}[a]$ [Q]

   (c) Une réflexion h de $\mathbb{R}^n$ est déterminée de manière unique par l'hyperplan (Π) par rapport auquel on effectue la symétrie orthogonale.

   Si la réflexion h existe, alors le vecteur non nul $u - v$ est orthogonal à l'hyperplan (Π), ce qui détermine (Π) et donc h de manière unique.

   Réciproquement, soit h la réflexion par rapport à l'hyperplan orthogonal à $u - v$.

   $$h(u) = u - \frac{2}{\|u-v\|^2} < u - v, u > (u - v) = u - \frac{1}{\|u-v\|^2} < u - v, 2u > (u - v)$$

   Or u et v ont même norme. Donc $< u - v, u + v >= \|u\|^2 - \|v\|^2 = 0$.

On peut alors écrire $< u - v, 2u > = < u - v, (u + v) + (u - v) > = \|u - v\|^2$.

On en déduit : $h(u) = u - \dfrac{1}{\|u - v\|^2}\,\|u - v\|^2\,(u - v) = v$.

Conclusion : si u et v sont distincts et de même norme, il existe une réflexion h et une seule qui transforme u en v.

Remarque : si $u = v$, toute réflextion h par rapport à un hyperplan contenant u laisse le vecteur u invariant et vérifie donc $h(u) = v \ldots$ [Q]

(d) Si u est lui-même colinéaire à e_1, toute réflexion laissant invariant e_1 convient.

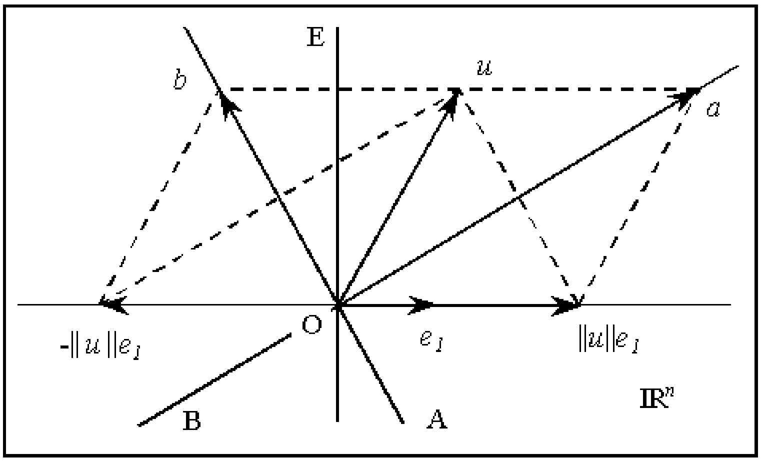
On suppose donc qu'il n'en est pas ainsi.

Puisqu'une réflexion conserve la norme, on a nécessairement $|\lambda| = \|u\|$. On a donc deux possibilités : $\lambda = \varepsilon\,\|u\|$, avec $\varepsilon \in \{-1, 1\}$.

Les vecteurs u et $v = \lambda e_1$ étant distincts, il y a une solution h unique pour chaque valeur de λ : c'est la réflexion h associée au vecteur $a = u - v = u - \varepsilon\,\|u\|\,e_1$.

Si $u = (x_1, x_2, \ldots, x_n)$, alors $a = (x_1 - \varepsilon\,\|u\|\,, x_2, \ldots, x_n)$.

La figure ci-dessous illustre la situation, avec des notations évidentes :



[Q]

(e) Si u est combinaison linéaire de $e_1, \ldots, e_j$, toute réflexion laissant invariant ces vecteurs convient. On suppose que tel n'est pas le cas, ce qui revient à dire que le vecteur w n'est pas colinéaire à e_j.

Avec les notations de l'énoncé, $\|v\|^2 = \displaystyle\sum_{k=1}^{j-1} x_k^2 + \lambda^2$ et $\|u\|^2 = \displaystyle\sum_{k=1}^{j-1} x_k^2 + \|w\|^2$.

On a donc nécessairement $|\lambda| = \|w\|$, c'est-à-dire $\lambda = \varepsilon\,\|w\|$, avec $\varepsilon \in \{-1, 1\}$.

Par hypothèse, u et v sont distincts. Pour chacune des deux valeurs de λ, il y a donc une solution unique qui est la réflexion h par rapport à l'hyperplan orthogonal à $a = u - v = (0\ldots, 0, u_j - \varepsilon\,\|w\|\,, u_{j+1}, \ldots, u_n) = w - \varepsilon\,\|w\|\,e_j$.

Puisque $e_1, \ldots, e_{j-1}$ sont orthogonaux au vecteur a, ils sont invariants par h. [Q]

2. (a) Dans cette question, on identifie les colonnes de A (et des matrices qui vont se déduire de A) avec des vecteurs de $\mathbb{R}^n$, et les matrices de Householder avec les transformations orthogonales correspondantes dans la base canonique de $\mathbb{R}^n$.
On va démontrer le résultat par une récurrence finie sur l'indice de colonne j.
   - Traitement de la première colonne :
     Soit u le premier vecteur-colonne de A.
     Si ce vecteur est colinéaire à e_1 (c'est-à-dire si les coefficients sous-diagonaux de la première colonne de A sont déjà nuls) on pose $H = I_n$.
     Sinon on sait qu'il existe une réflexion H_1 telle que le vecteur $v = H_1 u$ soit proportionnel à e_1. Mais v est le premier vecteur-colonne de $A_1 = H_1 A$, ce qui établit la propriété au rang 1.
   - Traitement de la j-ième colonne :
     Soit j un entier compris entre 1 et $n-2$. On suppose que les matrices $H_1, \ldots, H_{j-1}$ sont connues. Avec les notations de l'énoncé cela signifie que les coefficients sous-diagonaux des $j-1$ premières colonnes de A_{j-1} sont nuls.
     S'il en est de même pour le j-ème vecteur-colonne de A_{j-1}, alors on pose $H_j = I_n$ et donc $A_j = A_{j-1}$.
     Sinon soit u le j-ième vecteur-colonne de A_{j-1} : on sait qu'il existe une réflexion H_j qui envoie u sur un vecteur v combinaison linéaire de $e_1, \ldots, e_j$.
     On sait également que cette réflexion laisse invariant les vecteurs $e_1, \ldots, e_{j-1}$. Elle ne modifie donc pas les $j-1$ premiers vecteurs-colonnes de A_{j-1}.
     Ainsi les coefficients sous-diagonaux des j premières colonnes de $A_j = H_j A_{j-1}$ sont égaux à 0, ce qui démontre la propriété au rang j.
   - Conclusion :
     On a démontré l'existence des $n-1$ matrices de Householder $H_1, \ldots, H_{n-1}$.
     On peut remarquer qu'il n'y a pas unicité car à chaque étape on a en général deux possibilités suivant le signe donné au j-ème coefficient diagonal de A_j.

   [Q]

   (b) Avec les notations précédentes, la matrice $R = A_{n-1} = H_{n-1} \cdots H_2 H_1 A$ est triangulaire supérieure. Tout comme les matrices H_j et la matrice A, elle est inversible (ses coefficients diagonaux sont non nuls.)
On note que chaqueH_j est orthogonale et vérifie $H_j^{-1} = H_j$ (c'est normal pour des matrices de réflexion, mais c'est bien sûr vrai si certaines H_j sont égales à I_n.).
On peut alors écrire, en posant $Q = H_1 H_2 \cdots H_{n-1}$:

$$A = (H_{n-1} \cdots H_2 H_1)^{-1} = H_1^{-1} H_2^{-1} \cdots H_{n-1}^{-1} R = H_1 H_2 \cdots H_{n-1} R = QR$$

La matrice Q est orthogonale car elle est un produit de matrices orthogonales. On a donc obtenu une décomposition QR de la matrice A.
Dans la pratique, on passe de A à R en multipliant à gauche (et successivement) par la matrice H_1, puis par la matrice H_2, etc.
Si on applique ces mêmes multiplications mais à droite et en partant de I_n, alors on forme successivement les matrices H_1, $H_1 H_2$ et finalement $Q = H_1 H_2 \cdots H_{n-1}$. [Q]

(c) – Traitement de la première colonne :
Le premier vecteur colonne de A est $u_1 = (-4, -4, -8, 23)$.
La norme de ce vecteur est $\|u_1\| = \sqrt{4^2 + 4^2 + 8^2 + 23^2} = 25$.
On va transformer u_1 en le vecteur $v_1 = (25, 0, 0, 0)$ au moyen de la réflexion h_1 par rapport à l'hyperplan orthogonal au vecteur $a_1 = u_1 - \|u_1\| \, e_1 = (-29, -4, -8, 23)$.
Le carré de la norme de a_1 est :

$$2\|u_1\|^2 - 2\|u_1\| < u_1, e_1 >= 2\|u_1\|^2 + 8\|u_1\| = 2 \cdot 25 \cdot 29 = 2 \cdot 725$$

La matrice de Householder associée est :

$$H_1 = I_n - \frac{2}{\|a_1\|^2}[a_1]^{\mathrm{T}}[a_1] = I_n - \frac{1}{725}\begin{pmatrix} -29 \\ -4 \\ -8 \\ 23 \end{pmatrix}\begin{pmatrix} -29 & -4 & -8 & 23 \end{pmatrix}$$

$$= I_n - \frac{1}{725}\begin{pmatrix} 841 & 116 & 232 & -667 \\ 116 & 16 & 32 & -92 \\ 232 & 32 & 64 & -184 \\ -667 & -92 & -184 & 529 \end{pmatrix} = \frac{1}{725}\begin{pmatrix} -116 & -116 & -232 & 667 \\ -116 & 709 & -32 & 92 \\ -232 & -32 & 661 & 184 \\ 667 & 92 & 184 & 196 \end{pmatrix}$$

On en déduit :

$$A_1 = H_1 A = \frac{1}{725}\begin{pmatrix} -116 & -116 & -232 & 667 \\ -116 & 709 & -32 & 92 \\ -232 & -32 & 661 & 184 \\ 667 & 92 & 184 & 196 \end{pmatrix}\begin{pmatrix} -4 & 20 & 35 & 5 \\ -4 & -30 & -15 & 55 \\ -8 & 40 & -80 & -65 \\ 23 & -15 & 30 & 15 \end{pmatrix}$$

$$= \frac{25}{29}\begin{pmatrix} 29 & -29 & 58 & 29 \\ 0 & -42 & -15 & 167 \\ 0 & 32 & -88 & -69 \\ 0 & 24 & 21 & -1 \end{pmatrix}$$

– Traitement de la deuxième colonne :
Le deuxième vecteur colonne de A_1 est $u_2 = \dfrac{25}{29}(-29, -42, 32, 24)$.
Posons $w = \dfrac{25}{29}(0, -42, 32, 24)$ (voir les notations de la question II-1-e.)
La norme de ce vecteur est $\|w\| = \dfrac{25}{29}\sqrt{42^2 + 32^2 + 24^2} = 50$.
On sait qu'on doit appliquer la réflexion h_2 par rapport à l'hyperplan orthogonal à $w - \|w\| \, e_2 = \dfrac{100}{29}(0, -25, 8, 6)$. Pour simplifier, on choisit $a_2 = (0, -25, 8, 6)$.
Le carré de la norme de a_2 est : $25^2 + 8^2 + 6^2 = 725$.

La matrice de Householder associée est :

$$H_2 = I_n - \frac{2}{\|a_2\|^2}[a_2]^{\mathrm{T}}[a_2] = I_n - \frac{2}{725}\begin{pmatrix}0\\-25\\8\\6\end{pmatrix}(0,-25,8,6)$$

$$= I_n - \frac{2}{725}\begin{pmatrix}0&0&0&0\\0&625&-200&-150\\0&-200&64&48\\0&-150&48&36\end{pmatrix} = \frac{1}{725}\begin{pmatrix}725&0&0&0\\0&-525&400&300\\0&400&597&-96\\0&300&-96&653\end{pmatrix}$$

On en déduit la matrice $A_2 = H_2A_1 = H_2H_1A$:

$$A_2 = \frac{1}{725}\begin{pmatrix}725&0&0&0\\0&-525&400&300\\0&400&597&-96\\0&300&-96&653\end{pmatrix}\frac{25}{29}\begin{pmatrix}29&-29&58&29\\0&-42&-15&167\\0&32&-88&-69\\0&24&21&-1\end{pmatrix}$$

$$= \begin{pmatrix}25&-25&50&25\\0&50&-25&-75\\0&0&-72&-17\\0&0&21&31\end{pmatrix}$$

On trouve alors la matrice H_1H_2 :

$$H_1H_2 = \frac{1}{725}\begin{pmatrix}-116&-116&-232&667\\-116&709&-32&92\\-232&-32&661&184\\667&92&184&196\end{pmatrix}\frac{1}{725}\begin{pmatrix}725&0&0&0\\0&-525&400&300\\0&400&597&-96\\0&300&-96&653\end{pmatrix}$$

$$= \frac{1}{625}\begin{pmatrix}-100&200&-296&503\\-100&-425&304&328\\-200&400&433&56\\575&100&152&164\end{pmatrix}$$

– Traitement de la troisième colonne :

Le troisième vecteur colonne de A_2 est $u_3 = (50,-25,-72,21)$.

Posons $w = (0,0,-72,21)$. La norme de ce vecteur est $\|w\| = \sqrt{72^2+21^2} = 75$.

On sait qu'on doit appliquer la réflexion h_3 par rapport à l'hyperplan orthogonal à $a_3 = w - \|w\|\, e_3 = (0,0,-147,21)$.

Le carré de la norme de a_3 est : $147^2 + 21^2 = 22050$.

La matrice de Householder associée est :

$$H_3 = I_n - \frac{2}{\|a_3\|^2}\,[a_3]^{\mathrm{T}}[a_3] = I_n - \frac{1}{11025}\begin{pmatrix}0\\0\\-147\\21\end{pmatrix}(0,0,-147,21)$$

$$= I_n - \frac{1}{11025}\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&21609&-3087\\0&0&-3087&441\end{pmatrix} = \frac{1}{25}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-24&7\\0&0&7&24\end{pmatrix}$$

On en déduit la matrice $R = A_3 = H_3A_2 = H_3H_2H_1A$:

$$R = \frac{1}{25}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-24&7\\0&0&7&24\end{pmatrix}\begin{pmatrix}25&-25&50&25\\0&50&-25&-75\\0&0&-72&-17\\0&0&21&31\end{pmatrix} = \begin{pmatrix}25&-25&50&25\\0&50&-25&-75\\0&0&75&25\\0&0&0&25\end{pmatrix}$$

Enfin, voici $Q = H_1H_2H_3$:

$$Q = \frac{1}{625}\begin{pmatrix}-100&200&-296&503\\-100&-425&304&328\\-200&400&433&56\\575&100&152&164\end{pmatrix}\frac{1}{25}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-24&7\\0&0&7&24\end{pmatrix}$$

$$= \frac{1}{25}\begin{pmatrix}-4&8&17&16\\-4&-17&-8&16\\-8&16&-16&7\\23&4&-4&8\end{pmatrix}$$

La décomposition de A est alors terminée :

$$\underbrace{\begin{pmatrix}-4&20&35&5\\-4&-30&-15&55\\-8&40&-80&-65\\23&-15&30&15\end{pmatrix}}_{A\text{ inversible}} = \underbrace{\frac{1}{25}\begin{pmatrix}-4&8&17&16\\-4&-17&-8&16\\-8&16&-16&7\\23&4&-4&8\end{pmatrix}}_{Q\text{ orthogonale}}\underbrace{\begin{pmatrix}25&-25&50&25\\0&50&-25&-75\\0&0&75&25\\0&0&0&25\end{pmatrix}}_{R\text{ triangulaire}}$$

[Q]