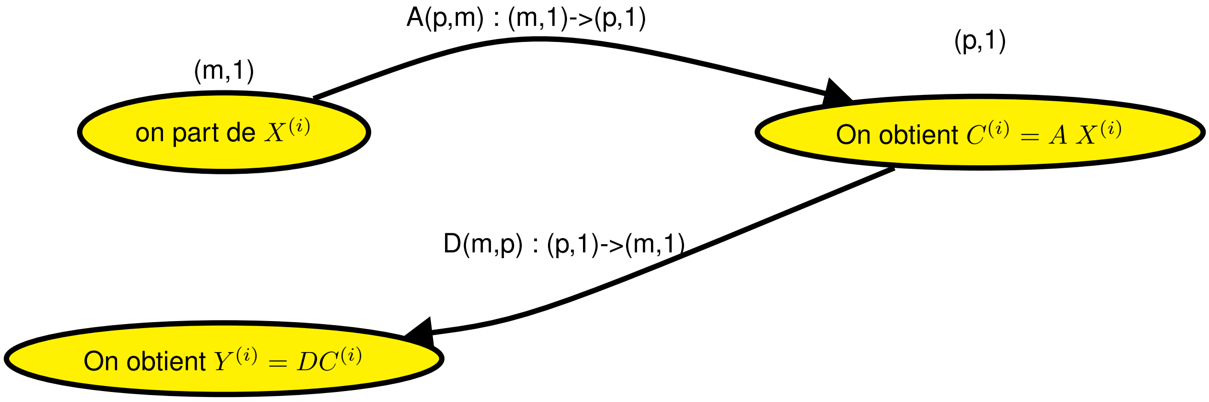
1 2.12 analyse en composantes principales : recherche de D et A optimums tels que d(X,DAX) minimum

On part d'un ou d'une série de vecteurs $(X^{(i)})_{(i=1..n)} = \begin{pmatrix} x_1^{(i)} \\ \vdots \\ x_m^{(i)} \end{pmatrix}$.

Éventuellement une base : $(b^{(i)})_{(i=1..m)} = \begin{pmatrix} b_1^{(i)} \\ \vdots \\ b_m^{(i)} \end{pmatrix}$. Tout X s'écrira $X = \sum_{i=1}^{i=m} x_i b^{(i)}$

On calcule l'image de $X^{(i)}$ vers un espace plus petit.



On cherche à minimiser $d^{(i)} = ||Y^{(i)} - X^{(i)}||^2$, qu'on note, d, Y et X et C.

$d = ||Y - X||^2 = ||DC - X||^2 = T(DC - X)(DC - X) = T(X)X - T(DC)X - T(X)DC + T(DC)DC = ||X||^2 - 2T(X)DC + T(DC)DC$

$$d = ||X||^2 - 2T(X)DC + T(C)T(D)DC$$

On suppose qu'il existe une solution optimale telle que

$$T(D)D = I_p$$

alors $d = ||X||^2 - 2T(X)DC + T(C)C$

On dérive selon l'axe $\varepsilon\ C$.

$$\nabla_C(d) = \lim_{\varepsilon \to 0} \frac{d(C+\varepsilon\ C) - d(C)}{\varepsilon} \tag{1}$$

$$\tag{2}$$

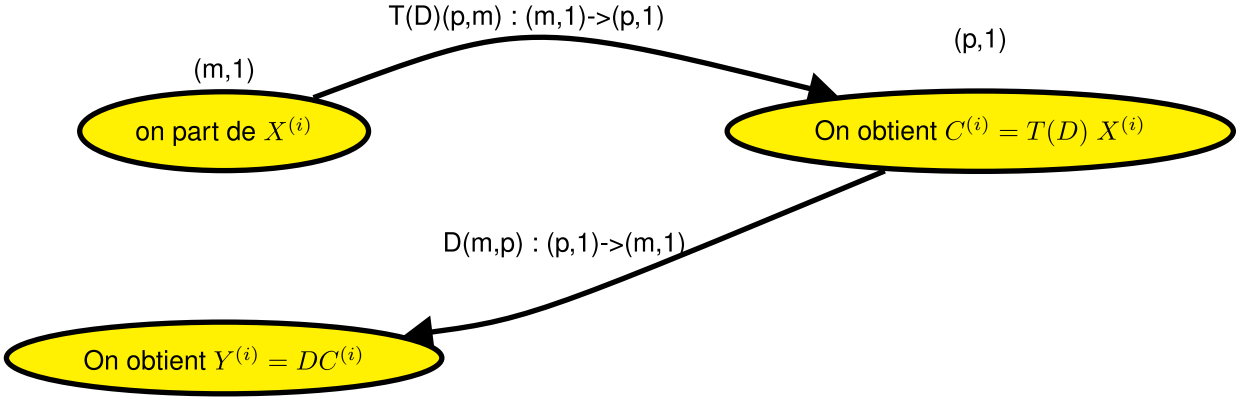
$$\begin{aligned} d(C+\varepsilon\ ) &= ||X||^2 - 2T(X)D(C+\varepsilon\ C) + T(C+\varepsilon\ C)(C+\varepsilon\ C) && (3) \\ &= ||X||^2 - 2T(X)DC - 2\varepsilon T(X)DC) + T(C)(C) + \varepsilon T(C)(C) + \varepsilon T(C)(C) + \varepsilon^2 T(C)(C) && (4) \end{aligned}$$

$$\begin{aligned} \nabla_C(d) &= \lim_{\varepsilon \to 0} \{-2T(X)DC + T(C)(C) + T(C)(C) + \varepsilon T(C)(C)\} && (5) \\ &= -2T(X)DC + 2T(C)(C) && (6) \end{aligned}$$

$$\nabla_C(d) = 0 \Leftrightarrow -2T(X)DC + 2T(C)(C) = 0 \Leftrightarrow 2(-T(X)D + T(C))C = 0 \Leftrightarrow ?T(C) = T(X)D \Leftrightarrow C = T(D)X \tag{7}$$

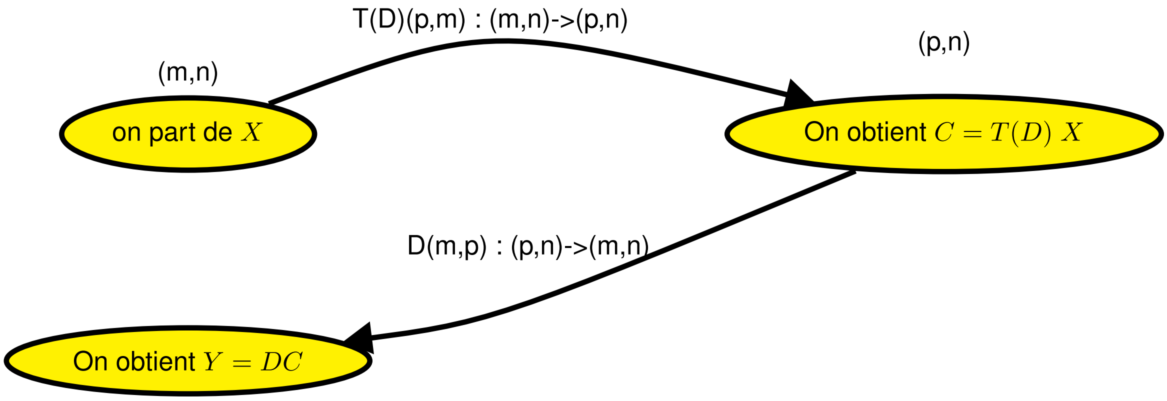
$$A = T(D)$$

On obtient :



(m,1)
on part de $X^{(i)}$
T(D)(p,m) : (m,1)->(p,1)
(p,1)
On obtient $C^{(i)} = T(D)\ X^{(i)}$
D(m,p) : (p,1)->(m,1)
On obtient $Y^{(i)} = DC^{(i)}$

On crée une matrice X (m,n) formée des vecteurs $(X^{(i)})_{(i=1..n)}$ dans la colonne i. $X_{ai} = X_a^{(i)}$



On cherche à minimiser

$$\begin{aligned}
d^2 &= ||X - Y||_F^2 & (8)\\
&= ||X - D\ T(D)\ X||_F^2 & (9)\\
d^2 &= \sum_{b=1...n} ||X^{(b)} - Y^{(b)}||_F^2 & (10)\\
&= \sum_{b=1...n} ||X^{(b)} - D\ T(D)\ X^{(b)}||_F^2 & (11)\\
&= \sum_{a=1...m;b=1...n} (Y_{ab} - X_{ab})^2 = \sum_{a=1...m;b=1...n} \left(\sum_{e=1...p;f=1...m} D_{ae} D_{fe} X_{fb} - X_{ab} \right)^2 & (12)
\end{aligned}$$

—

1.1 propriété des traces

Soit A une matrice ; $d^2 = ||A||_F^2 = \sum_{ij} A_{ij}^2$.
$A = A_{ij}$; $AT(A) = (\sum_k A_{ik} A_{jk})_{ij}$; $Tr(AT(A)) = \sum_i \sum_k A_{ik} A_{ik} = \sum_{i,k} A_{ik}^2$

$$||A||_F = \sqrt{Tr(AT(A))}$$

$Tr(A(B + C)) = \sum_i (A(B + C))_{ii} = \sum_{i,k} (A_{ik}(B_{ki} + C_{ki})) = \sum_{i,k} (A_{ik}(B_{ki})) + \sum_{i,k} (A_{ik} C_{ki}) = Tr(AB) + Tr(AC)$

—

$\left\{ \begin{array}{l} (D\ T(D)\ X^{(b)})_a = \left(\sum_{e=1...p;f=1...m} D_{ae} D_{fe} X_f \right)_a \\ (T(X^{(b)}) D\ T(D))_a = \left(\sum_{f=1...m;e=1...p} X_f D_{fe} D_{ae} \right)_a \end{array} \right. \Rightarrow D\ T(D)\ X^{(b)} = T(X^{(b)}) D\ T(D) \Rightarrow d^2 = \sum_{b=1...n} ||X^{(b)} - X^{(b)}\ D\ T(D)||_F^2$

$$d^2 = ||X - X\ D\ T(D)||_F^2$$

$$\begin{aligned}
d^2 &= ||X - X\ D\ T(D)||_F^2 & (13)\\
&= Tr(X - X\ D\ T(D))T(X - X\ D\ T(D))) & (14)\\
&= Tr(XT(X) + XT(-X\ D\ T(D)) - X\ D\ T(D)T(X) - X\ D\ T(D)T(-X\ D\ T(D))) & (15)\\
&= Tr(XT(X)) + Tr(XT(-X\ D\ T(D)) - Tr(X\ D\ T(D)T(X) + Tr(X\ D\ T(D)T(X\ D\ T(D))) & (16)\\
&= Tr(XT(X)) - 2Tr(XD\ T(D)T(X)) + Tr(X\ D\ T(D)DT(D)T(X))) & (17)\\
&= Tr(XT(X)) - 2Tr(T(X)XD\ T(D)) + Tr(T(X)X\ D\ T(D)DT(D))) & (18)\\
&= Tr(XT(X)) - 2Tr(T(X)XD\ T(D)) + Tr(T(X)X\ D\ T(D))) & (19)
\end{aligned}$$

$$d^2 = Tr(XT(X)) - Tr(T(D)T(X)XD\ )$$

X étant fixe, on cherche à minimiser d donc à maximiser $Tr(T(D)T(X)XD\ )$ $T(X)X$ est une matrice (n,n) éventuellement diagonalisable dans la base (G_i, λ_i). donc $T(X)X = Q^{-1} Diag(\lambda_i) Q$

1.2 résolution avec $X^{(1)}$ et $X^{(2)}$, dim 2 => dim 1

On part d'un ou d'une série de vecteurs $X^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ et $X^{(2)} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$

$X = (X_{i,j})\,;\, X_{i,j} = X^{(i)}_j = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$

$D = \begin{pmatrix} a \\ b \end{pmatrix}$ donc $T(D) = \begin{pmatrix} a & b \end{pmatrix}$

$T(D)D = I <=> \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = I <=> a^2 + b^2 = 1$

$T(D)X^{(1)} = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = (a + 2b)\,;\, T(D)X^{(2)} = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix} = (4a + 5b)$

$DT(D)X^{(1)} = \begin{pmatrix} a \\ b \end{pmatrix} (a + 2b) = \begin{pmatrix} a(a + 2b) \\ b(a + 2b) \end{pmatrix}\,;\, DT(D)X^{(2)} = \begin{pmatrix} a \\ b \end{pmatrix} (4a + 5b) = \begin{pmatrix} a(4a + 5b) \\ b(4a + 5b) \end{pmatrix}$

$a = cos\theta; b = sin\theta$

$$\begin{aligned}
d^2 &= ||X^{(1)} - DT(D)X^{(1)}||_F^2 + ||X^{(2)} - DT(D)X^{(2)}||_F^2 && (20)\\
&= (1 - a(a + 2b))^2 + (2 - b(a + 2b))^2 + (4 - a(4a + 5b))^2 + (5 - b(4a + 5b))^2 && (21)\\
&= 1 + 4 + 16 + 25 - 2(a(a + 2b) + b(a + 2b) + a(4a + 5b) + b(4a + 5b)) && (22)\\
&\quad + a^2(a + 2b)^2 + b^2(a + 2b)^2 + a^2(4a + 5b)^2 + b^2(4a + 5b)^2 && (23)\\
&= 46 - 2(5a^2 + 12ab + 7b^2) + (17a^4 + 46a^2b^2 + 29b^4 + 44a^3b + 44ab^3) && (24)\\
&= 46 - 2(5a^2 + 12a\sqrt{1 - a^2} + 7 - 7a^2) + && (25)\\
&\quad (17a^4 + 46a^2(1 - a^2) + 29(1 - a^2)^2 + 44a^3\sqrt{1 - a^2} + 44a\sqrt{1 - a^2}(1 - a^2)) && (26)\\
&= (46 - 7 + 29) + a^2(-10 + 14 + 46 + 29(-2)) + 12a\sqrt{1 - a^2} + a^4(17 - 46 + 29) + 44a\sqrt{1 - a^2} && (27)\\
&= 68 - 8a^2 + 56a\sqrt{1 - a^2} && (28)\\
& && (29)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial d^2}{\partial a} &= -16a + 56\sqrt{1 - a^2} + 56a\frac{-2a}{\sqrt{1 - a^2}} = \frac{16a(1 - a^2) + 56(1 - a^2)\sqrt{1 - a^2} + 56a(-2a)\sqrt{1 - a^2}}{1 - a^2} && (30)\\
\frac{\partial d^2}{\partial a} = 0 &\Leftrightarrow 56(1 - 3a^2)\sqrt{1 - a^2} = 16a(a^2 - 1) \Rightarrow 56^2(1 - 3a^2)^2(1 - a^2) = 16^2a^2(a^2 - 1)^2 && (31)\\
&\quad 56^2(1 - 3a^2)^2 = 16^2a^2(a^2 - 1) \Rightarrow 27968a^4 - 18800a^2 + 3136 = 0; \delta = 2609408; && (32)
\end{aligned}$$

On cherche à minimiser $d^{(i)} = ||Y^{(i)} - X^{(i)}||^2$, qu'on note, d, Y et X et C.

$d = ||Y - X||^2 = ||DC - X||^2 = T(DC - X)(DC - X) = T(X)X - T(DC)X - T(X)DC + T(DC)DC = ||X||^2 - 2T(X)DC + T(DC)DC$

$$d = ||X||^2 - 2T(X)DC + T(C)T(D)DC$$

On suppose qu'il existe une solution optimale telle que

$$T(D)D = I_p$$

alors $d = ||X||^2 - 2T(X)DC + T(C)C$

On dérive selon l'axe $\varepsilon\ C$.

$$\begin{aligned} \nabla_C(d) &= \lim_{\varepsilon \to 0} \tfrac{d(C+\varepsilon\ C) - d(C)}{\varepsilon} && (33) \\ & && (34) \end{aligned}$$

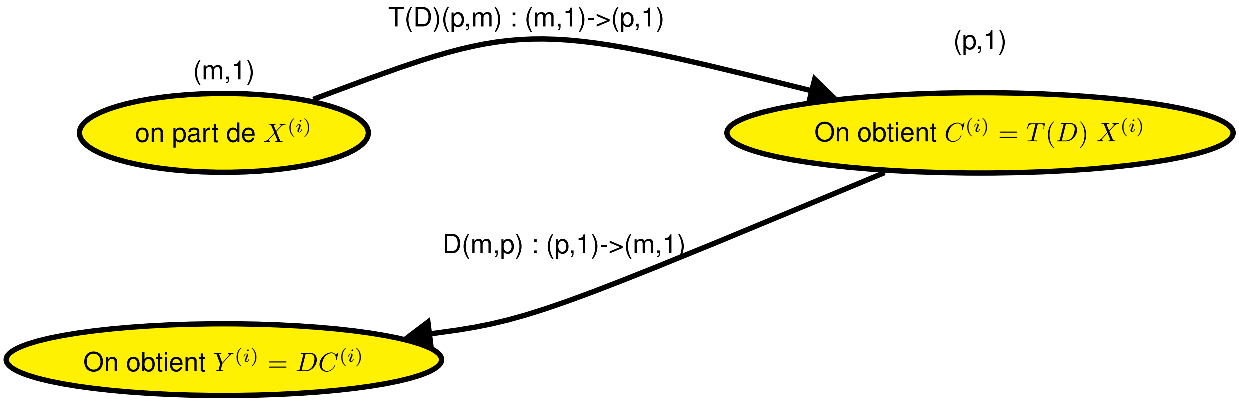
$$\begin{aligned} d(C+\varepsilon\ ) &= ||X||^2 - 2T(X)D(C+\varepsilon\ C) + T(C+\varepsilon\ C)(C+\varepsilon\ C) && (35) \\ &= ||X||^2 - 2T(X)DC - 2\varepsilon T(X)DC) + T(C)(C) + \varepsilon T(C)(C) + \varepsilon T(C)(C) + \varepsilon^2 T(C)(C) && (36) \end{aligned}$$

$$\begin{aligned} \nabla_C(d) &= \lim_{\varepsilon \to 0} \{-2T(X)DC + T(C)(C) + T(C)(C) + \varepsilon T(C)(C)\} && (37) \\ &= -2T(X)DC + 2T(C)(C) && (38) \end{aligned}$$

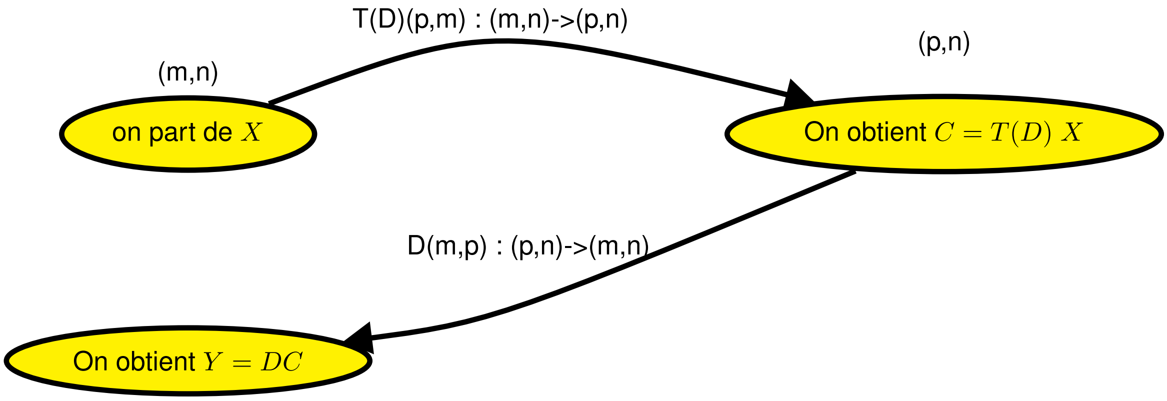
$$\nabla_C(d) = 0 \Leftrightarrow -2T(X)DC + 2T(C)(C) = 0 \Leftrightarrow 2(-T(X)D + T(C))C = 0 \Leftrightarrow ?T(C) = T(X)D \Leftrightarrow C = T(D)X \tag{39}$$

$$A = T(D)$$

On obtient :



On crée une matrice X (m,n) formée des vecteurs $(X^{(i)})_{(i=1..n)}$ dans la colonne i. $X_{ai} = X_a^{(i)}$



On cherche à minimiser

$$\begin{aligned}
d^2 &= ||X - Y||_F^2 &(40)\\
&= ||X - D\ T(D)\ X||_F^2 &(41)\\
d^2 &= \sum_{b=1\dots n} ||X^{(b)} - Y^{(b)}||_F^2 &(42)\\
&= \sum_{b=1\dots n} ||X^{(b)} - D\ T(D)\ X^{(b)}||_F^2 &(43)\\
&= \sum_{a=1\dots m; b=1\dots n} (Y_{ab} - X_{ab})^2 = \sum_{a=1\dots m; b=1\dots n} \left(\sum_{e=1\dots p; f=1\dots m} D_{ae} D_{fe} X_{fb} - X_{ab} \right)^2 &(44)
\end{aligned}$$

—

1.3 propriété des traces

Soit A une matrice ; $d^2 = ||A||_F^2 = \sum_{ij} A_{ij}^2$.
$A = A_{ij}$; $AT(A) = (\sum_k A_{ik} A_{jk})_{ij}$; $Tr(AT(A)) = \sum_i \sum_k A_{ik} A_{ik} = \sum_{i,k} A_{ik}^2$

$$||A||_F = \sqrt{Tr(AT(A))}$$

$Tr(A(B+C)) = \sum_i (A(B+C))_{ii} = \sum_{i,k} (A_{ik}(B_{ki} + C_{ki})) = \sum_{i,k}(A_{ik}(B_{ki})) + \sum_{i,k}(A_{ik} C_{ki})) = Tr(AB) + Tr(AC)$

—

$$\left\{ \begin{array}{l} (D\ T(D)\ X^{(b)})_a = \left(\sum_{e=1\dots p; f=1\dots m} D_{ae} D_{fe} X_f \right)_a \\ (T(X^{(b)}) D\ T(D))_a = \left(\sum_{f=1\dots m; e=1\dots p} X_f D_{fe} D_{ae} \right)_a \end{array} \right. \Rightarrow D\ T(D)\ X^{(b)} = T(X^{(b)}) D\ T(D) \Rightarrow d^2 = \sum_{b=1\dots n} ||X^{(b)} - X^{(b)}\ D\ T(D)||_F^2$$

$$d^2 = ||X - X\ D\ T(D)||_F^2$$

$$\begin{aligned}
d^2 &= ||X - X\ D\ T(D)||_F^2 &(45)\\
&= Tr(X - X\ D\ T(D)) T(X - X\ D\ T(D))) &(46)\\
&= Tr(XT(X) + XT(-X\ D\ T(D)) - X\ D\ T(D) T(X) - X\ D\ T(D) T(-X\ D\ T(D))) &(47)\\
&= Tr(XT(X)) + Tr(XT(-X\ D\ T(D)) - Tr(X\ D\ T(D) T(X) + Tr(X\ D\ T(D) T(X\ D\ T(D))) &(48)\\
&= Tr(XT(X)) - 2Tr(XD\ T(D) T(X)) + Tr(X\ D\ T(D) DT(D) T(X))) &(49)\\
&= Tr(XT(X)) - 2Tr(T(X) XD\ T(D)) + Tr(T(X) X\ D\ T(D) DT(D))) &(50)\\
&= Tr(XT(X)) - 2Tr(T(X) XD\ T(D)) + Tr(T(X) X\ D\ T(D))) &(51)
\end{aligned}$$

$$d^2 = Tr(XT(X)) - Tr(T(D) T(X) XD\ )$$

X étant fixe, on cherche à minimiser d donc à maximiser $Tr(T(D)T(X)XD\ )$ $T(X)X$ est une matrice (n,n) éventuellement diagonalisable dans la base (G_i, λ_i). donc $T(X)X = Q^{-1} Diag(\lambda_i) Q$