

## 2.1. Présentation de l'exemple

**/40**

	Stat	Math	Cpta	G° Fi	<i>Moyenne</i>
Individu n° 1 +	19	14	8	18	11.50
Individu n° 2 -	20	12	4	4	9.00
Individu n° 3 +	10	10	32	38	13.75
Individu n° 4 -	13	17	4	4	8.50
Individu n° 5 -	6	8	26	24	9.75
Individu n° 6 -	6	3	28	32	9.75
Individu n° 7 +	19	16	8	20	12.25
Individu n° 8 -	15	18	6	6	9.75
Individu n° 9 +	9	2	32	30	10.50
Individu n° 10 -	8	7	20	20	8.75
<i>Moyenne</i>	12.50	10.70	16.80	19.60	
<i>Ecart-type</i>	5.20	5.39	11.32	11.38	

## 2.2. Méthodologie de calcul

### 2.2.1. Rappels statistiques.

#### 1. Moyenne (Espérance), Médiane

En 2013, le salaire mensuel en équivalent temps plein (EQTP) dans le secteur privé ou dans les entreprises publiques est, en moyenne, de 2 912 euros bruts et de 2 202 euros nets des prélèvements à la source.

Le salaire médian, lui (qui sépare la population salariée en deux, la moitié gagnant plus, l'autre moins) s'établit à 1 772 euros net.

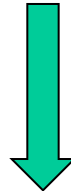
Source : [Insee](#)

Figure 3 - Distribution des salaires mensuels nets en 2013 et évolution entre 2012 et 2013 en euros constants

Déciles	Ensemble		Hommes		Femmes	
	2013	Évolution (%)	2013	Évolution (%)	2013	Évolution (%)
D1	1 200	-0,6	1 254	-0,9	1 154	-0,3
D2	1 342	-0,5	1 415	-0,6	1 268	-0,3
D3	1 471	-0,4	1 559	-0,5	1 374	-0,2
D4	1 609	-0,3	1 709	-0,3	1 485	-0,1
<b>D5 ou Médiane</b>	<b>1 772</b>	<b>-0,1</b>	<b>1 882</b>	<b>-0,2</b>	<b>1 619</b>	<b>0,1</b>
D6	1 974	0,0	2 100	-0,1	1 794	0,3
D7	2 244	0,0	2 405	0,0	2 029	0,3
D8	2 682	0,0	2 921	-0,1	2 368	0,3
D9	3 544	-0,2	3 892	-0,1	3 036	0,0
C95	4 526	-0,2	5 030	-0,4	3 756	0,2
C99	8 061	-0,5	9 253	-0,6	6 053	0,1
<b>Moyenne</b>	<b>2 202</b>	<b>-0,3</b>	<b>2 389</b>	<b>-0,4</b>	<b>1 934</b>	<b>0,0</b>

#### 2. Variance, Ecart-type

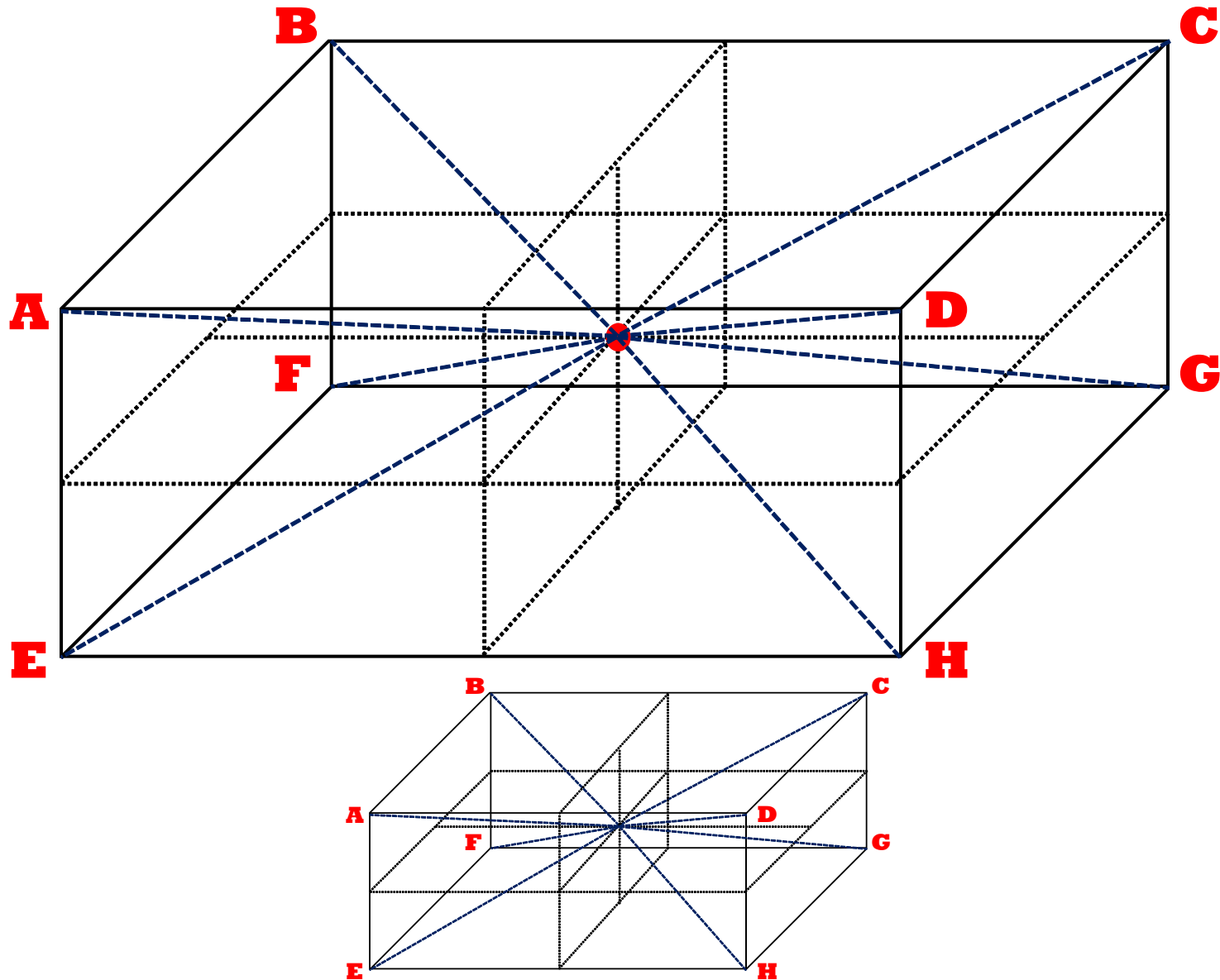
#### 3. Covariance, [coefficient de corrélation linéaire](#)



Effet d'une transformation linéaire

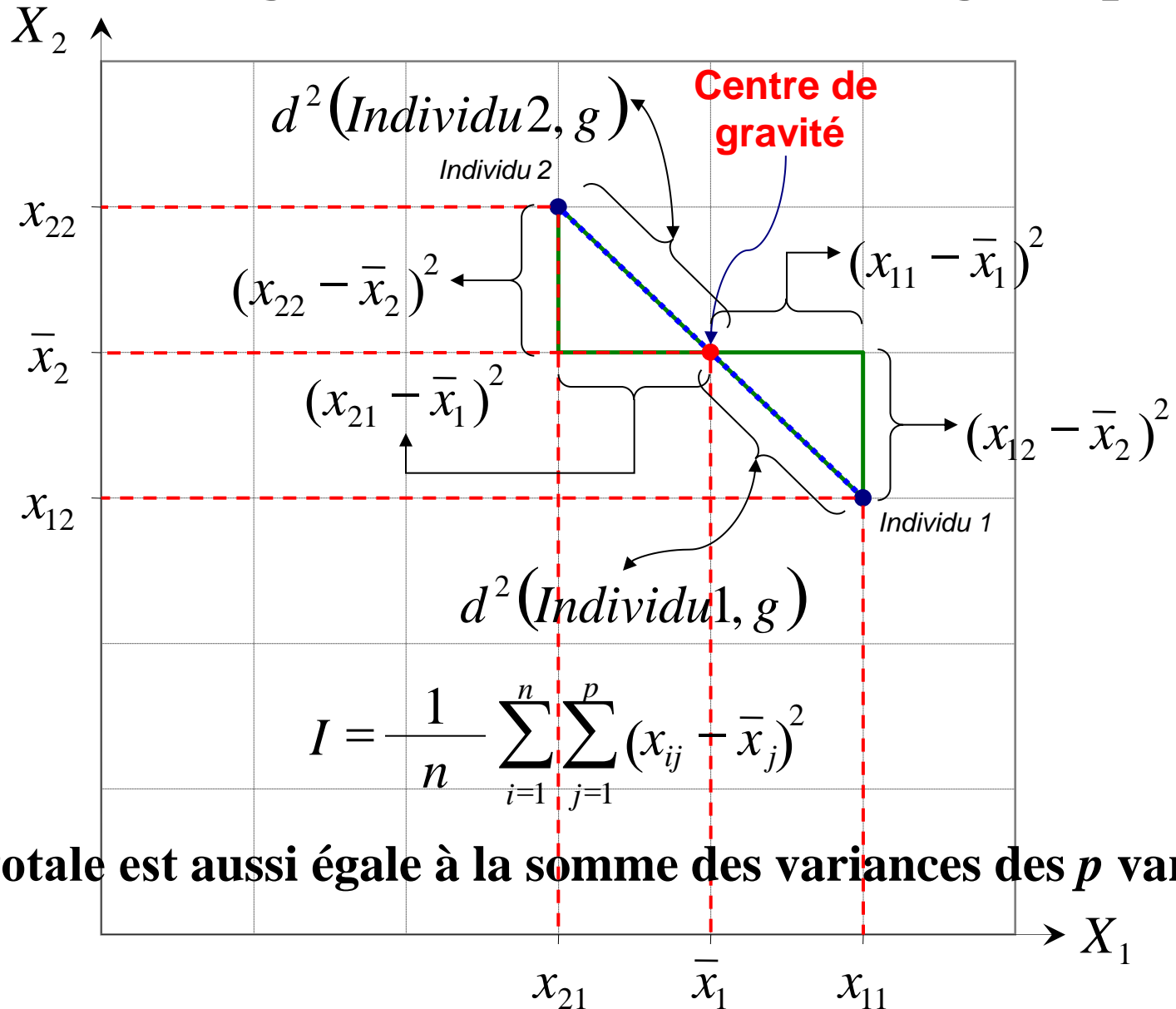
## 2.2. Méthodologie de calcul

### 2.2.2. Centre de gravité et inertie d'un nuage de points.



## 2.2. Méthodologie de calcul

### 2.2.2. Centre de gravité et inertie d'un nuage de points.



L'inertie totale est aussi égale à la somme des variances des  $p$  variables

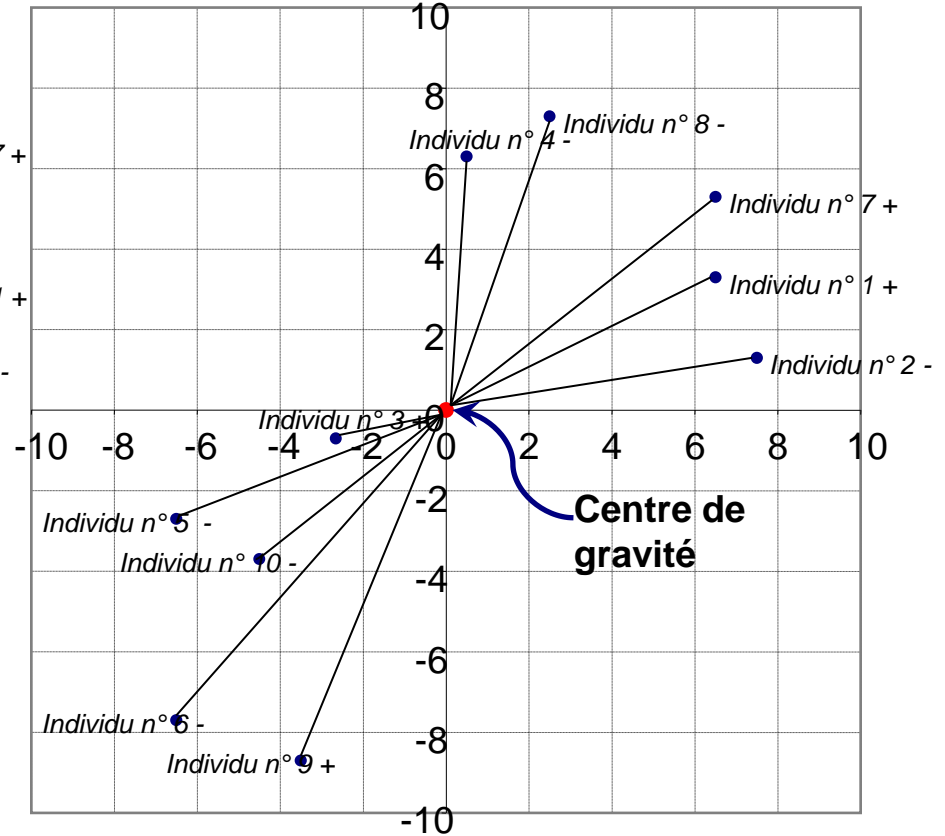
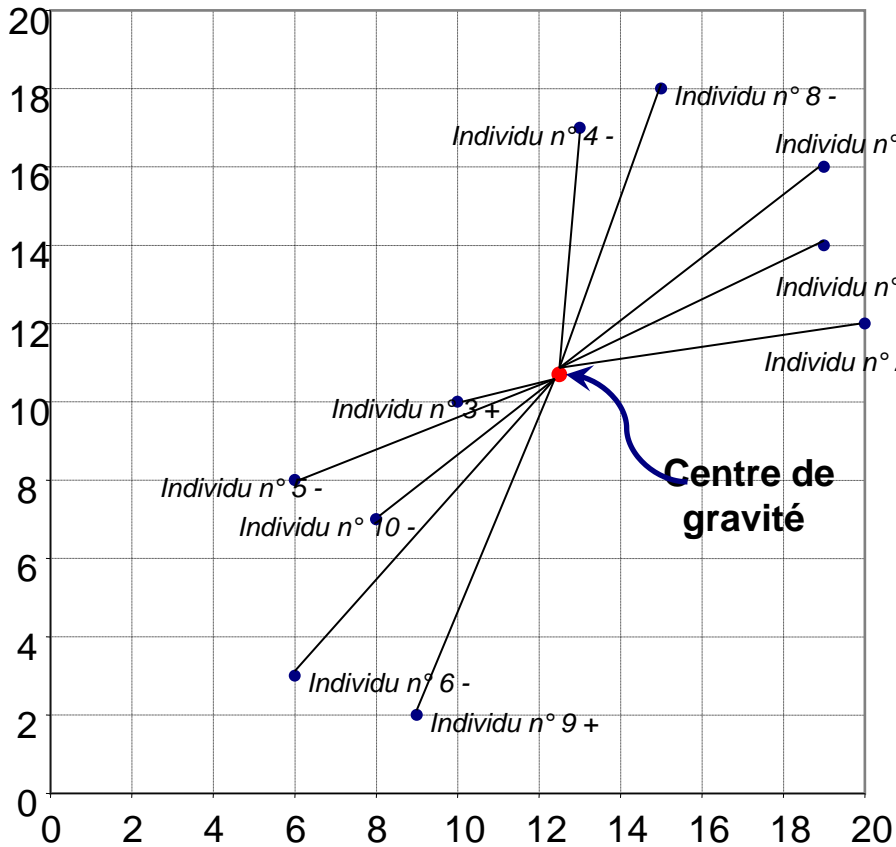
## 2.2. Méthodologie de calcul

### 2.2.3. Centrage et réduction des variables.

	Stat	Math	Cpta	G° Fi	<i>Moyenne</i>
Individu n° 1 +	6.50	3.30	-8.80	-1.60	1.15
Individu n° 2 -	7.50	1.30	-12.80	-15.60	-1.35
Individu n° 3 +	-2.50	-0.70	15.20	18.40	3.40
Individu n° 4 -	0.50	6.30	-12.80	-15.60	-1.85
Individu n° 5 -	-6.50	-2.70	9.20	4.40	-0.60
Individu n° 6 -	-6.50	-7.70	11.20	12.40	-0.60
Individu n° 7 +	6.50	5.30	-8.80	0.40	1.90
Individu n° 8 -	2.50	7.30	-10.80	-13.60	-0.60
Individu n° 9 +	-3.50	-8.70	15.20	10.40	0.15
Individu n° 10 -	-4.50	-3.70	3.20	0.40	-1.60
<i>Moyenne</i>	0.00	0.00	0.00	0.00	
<i>Ecart-type</i>	5.20	5.39	11.32	11.38	

# 2.2. Méthodologie de calcul

## 2.2.3. Centrage et réduction des variables.



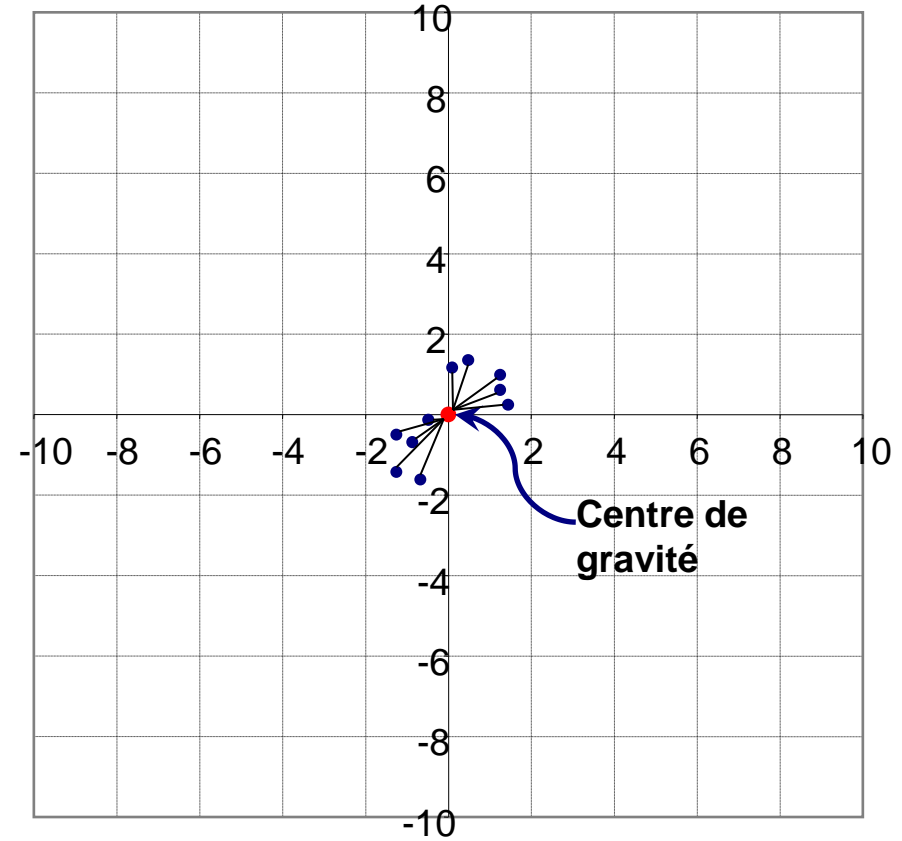
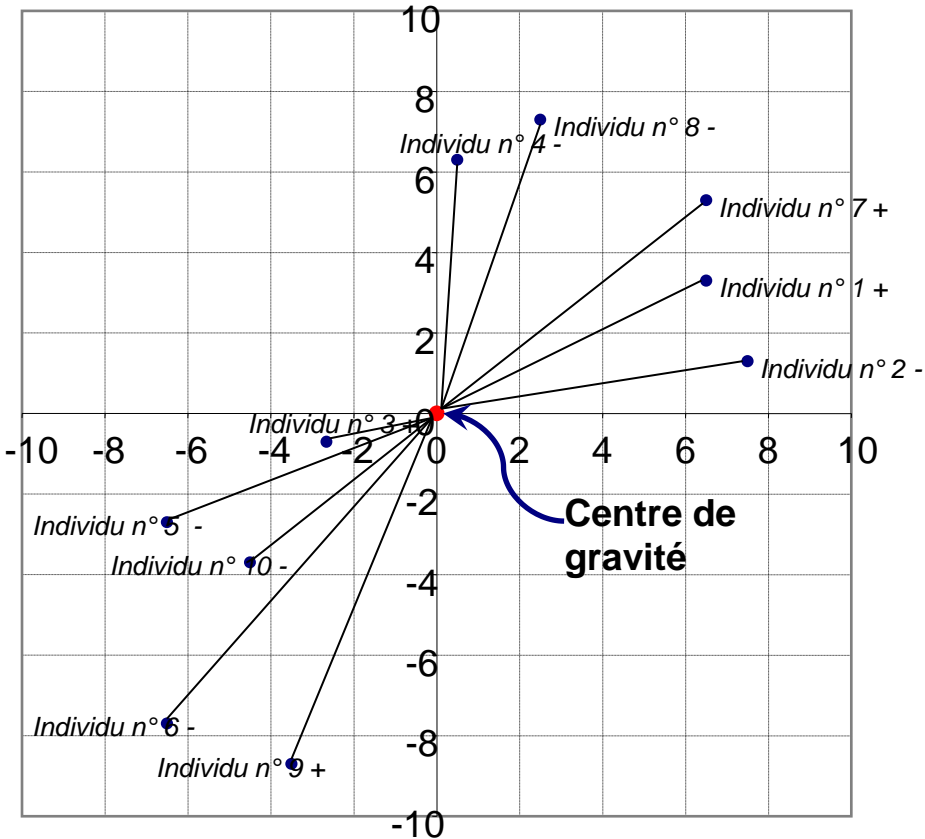
## 2.2. Méthodologie de calcul

### 2.2.3. Centrage et réduction des variables.

	Statistiques	Math	Cpta	G° Fi
Ind1+	1.2498	0.6127	-0.7773	-0.1406
Ind2-	1.4420	0.2414	-1.1307	-1.3712
Ind3+	-0.4807	-0.1300	1.3427	1.6173
Ind4-	0.0961	1.1697	-1.1307	-1.3712
Ind5-	-1.2498	-0.5013	0.8127	0.3867
Ind6-	-1.2498	-1.4296	0.9893	1.0899
Ind7+	1.2498	0.9840	-0.7773	0.0352
Ind8-	0.4807	1.3553	-0.9540	-1.1954
Ind9+	-0.6730	-1.6153	1.3427	0.9141
Ind10-	-0.8652	-0.6870	0.2827	0.0352
Moyenne	0	0	0	0
Ecart-type	1	1	1	1

## 2.2. Méthodologie de calcul

### 2.2.3. Centrage et réduction des variables.

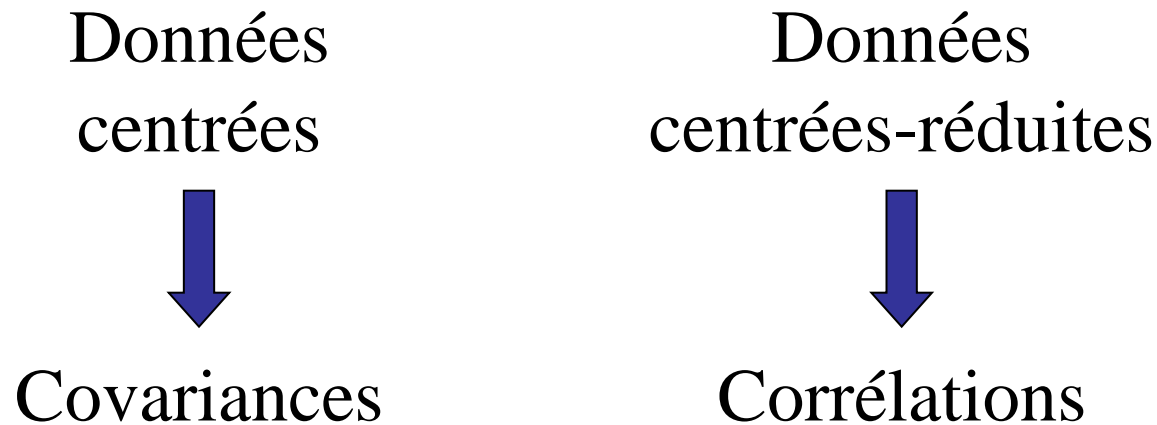




## 2.2. Méthodologie de calcul

### 2.2.4. Centrage et réduction des variables.

Dans une ACP les données sont systématiquement centrées





## 2.2. Méthodologie de calcul

### 2.2.5. Illustration de la logique de calcul.

$$c_1^2 + c_2^2 \neq 1$$

Cas	$c_1$	$c_2$	$Y_1$	$Var(Y_1)$
1	0.50	0.50	(-3.5, -1.0, 4.5)	11.17
2	0.50	-0.50	(-2.5, 1.0, 1.5)	3.17
3	1.58	-1.58	(-7.9, 3.2, 4.7)	31.43
4	10.00	-10.00	(-50, 20, 30)	1266.67
5	20.00	-20.00	(-100, 40, 60)	5066.67

$$c_1^2 + c_2^2 = 1$$

Cas	$c_1$	$c_2$	$Y_1$	$Var(Y_1)$
6	0.71	0.71	(-4.9, -1.4, 6.4)	22.33
7	0.71	-0.71	(-3.5, 1.4, 2.1)	6.33
8	0.94	-0.34	(-5.3, 0.7, 4.6)	16.68
9	0.94	0.34	(-6.0, -0.7, 6.7)	26.88

## 2.2. Méthodologie de calcul

### 2.2.5. Illustration de la logique de calcul.

1) si on ne fait que centrer les données, cette droite est engendrée par le **vecteur propre de la matrice des variances-covariances  $C$  associé à la plus grande valeur propre** (maximisation de l'inertie),

Cette logique de maximisation de l'inertie donne mécaniquement beaucoup d'importance aux variables les plus dispersées (unité de mesure)

2) si on centre et on réduit les données, cette droite est engendrée par le **vecteur propre de la matrice des corrélations  $R$  associé à la plus grande valeur propre** (maximisation de l'inertie).