

Annexe A

Révision de certains concepts de statistiques

A.1 Quelques distributions bien connues

Loi Normale : Si $Y \sim N(\mu, \sigma^2)$, sa densité est

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(y - \mu)^2}{2\sigma^2} \right\}.$$

Si $Z = \frac{Y - \mu}{\sigma}$, alors on dit que Z suit une loi normale centrée réduite, notée $N(0, 1)$.

Loi Khi-Carrée : Si $X \sim \chi^2_{(\nu)}$, sa densité est

$$f_X(x) = \frac{x^{\nu/2-1} e^{-x/2}}{2^{\nu/2} \Gamma(\nu/2)}, \quad x > 0.$$

On note que $E[X] = \nu$.

Si $Z \sim N(0, 1)$, alors $Z^2 \sim \chi^2_{(1)}$.

Si $Z_i \sim N(0, 1), i = 1, \dots, n$, alors

$$\sum_{i=1}^n Z_i^2 \sim \chi^2_{(n)}.$$

Loi Student t : Si $T \sim t_{(\nu)}$, sa densité est

$$f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}.$$

Quand $\nu \rightarrow \infty$, la loi t tend vers la loi Normale centrée réduite.

Si $Z \sim N(0, 1)$ et $X \sim \chi^2_{(\nu)}$, $Z \perp X$, alors

$$\frac{Z}{\sqrt{X/\nu}} \sim t_{(\nu)}.$$

Loi de Fisher : Si $X_1 \sim \chi^2_{(\nu_1)}$ et $X_2 \sim \chi^2_{(\nu_2)}$, $X_1 \perp X_2$, alors

$$\frac{X_1/\nu_1}{X_2/\nu_2} \sim F(\nu_1, \nu_2).$$

Aussi, si $T \sim t_{(\nu)}$, alors $T^2 \sim F(1, \nu)$.

A.2 Maximum de vraisemblance

Soit l'échantillon indépendant y_1, \dots, y_n , avec densités $f_{Y_i}(y_i; \theta)$, pour $i = 1, \dots, n$. La fonction de vraisemblance est

$$L(\theta; y_1, \dots, y_n) = \prod_{i=1}^n f_{Y_i}(y_i; \theta).$$

La fonction de log-vraisemblance est

$$l(\theta; y_1, \dots, y_n) = \ln L(\theta; y_1, \dots, y_n) = \sum_{i=1}^n \ln f_{Y_i}(y_i; \theta).$$

La méthode du maximum de vraisemblance permet de trouver l'estimateur $\hat{\theta}_n$ qui maximise $L(\theta; y_1, \dots, y_n)$, et par conséquent $l(\theta; y_1, \dots, y_n)$. On travaille habituellement avec le log pour simplifier les calculs. La fonction de score est

$$\dot{l}_{\theta}(\theta; y_1, \dots, y_n) = \frac{\partial}{\partial \theta} l(\theta; y_1, \dots, y_n).$$

L'estimateur du maximum de vraisemblance (EMV) est $\hat{\theta}_n$ tel que

$$\dot{l}_{\theta}(\theta; y_1, \dots, y_n) \Big|_{\hat{\theta}_n} = 0.$$

A.3 Estimateur non biaisé

Soit Y_1, \dots, Y_n , iid, avec densité $f(y; \theta)$. Soit l'estimateur de θ suivant : $\hat{\theta}_n = \hat{\theta}(Y_1, \dots, Y_n)$. On dit de $\hat{\theta}_n$ qu'il est non biaisé si

$$E[\hat{\theta}_n] = \theta.$$

Dans ce cas, le biais est zéro,

$$bias(\hat{\theta}_n) = E[\hat{\theta}_n] - \theta = 0,$$

ce qui réduit l'erreur quadratique moyenne (MSE) de l'estimateur :

$$MSE(\hat{\theta}_n) = E[(\hat{\theta}_n - \theta)^2] = Var(\hat{\theta}_n) + bias(\hat{\theta}_n).$$

A.4 Tables

Table de la distribution Student t

Les données dans le tableau sont $t_\nu(\alpha)$ tel que $P[|X| > t_\nu(\alpha)] = \alpha$ si $X \sim t_\nu$.
colonne = α ; rangée = ν

ν	0.900	0.700	0.500	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.158	0.510	1.000	1.963	3.078	6.314	12.706	25.452	63.657	636.619
2	0.142	0.445	0.816	1.386	1.886	2.920	4.303	6.205	9.925	31.599
3	0.137	0.424	0.765	1.250	1.638	2.353	3.182	4.177	5.841	12.924
4	0.134	0.414	0.741	1.190	1.533	2.132	2.776	3.495	4.604	8.610
5	0.132	0.408	0.727	1.156	1.476	2.015	2.571	3.163	4.032	6.869
6	0.131	0.404	0.718	1.134	1.440	1.943	2.447	2.969	3.707	5.959
7	0.130	0.402	0.711	1.119	1.415	1.895	2.365	2.841	3.499	5.408
8	0.130	0.399	0.706	1.108	1.397	1.860	2.306	2.752	3.355	5.041
9	0.129	0.398	0.703	1.100	1.383	1.833	2.262	2.685	3.250	4.781
10	0.129	0.397	0.700	1.093	1.372	1.812	2.228	2.634	3.169	4.587
11	0.129	0.396	0.697	1.088	1.363	1.796	2.201	2.593	3.106	4.437
12	0.128	0.395	0.695	1.083	1.356	1.782	2.179	2.560	3.055	4.318
13	0.128	0.394	0.694	1.079	1.350	1.771	2.160	2.533	3.012	4.221
14	0.128	0.393	0.692	1.076	1.345	1.761	2.145	2.510	2.977	4.140
15	0.128	0.393	0.691	1.074	1.341	1.753	2.131	2.490	2.947	4.073
16	0.128	0.392	0.690	1.071	1.337	1.746	2.120	2.473	2.921	4.015
17	0.128	0.392	0.689	1.069	1.333	1.740	2.110	2.458	2.898	3.965
18	0.127	0.392	0.688	1.067	1.330	1.734	2.101	2.445	2.878	3.922
19	0.127	0.391	0.688	1.066	1.328	1.729	2.093	2.433	2.861	3.883
20	0.127	0.391	0.687	1.064	1.325	1.725	2.086	2.423	2.845	3.850
21	0.127	0.391	0.686	1.063	1.323	1.721	2.080	2.414	2.831	3.819
22	0.127	0.390	0.686	1.061	1.321	1.717	2.074	2.405	2.819	3.792
23	0.127	0.390	0.685	1.060	1.319	1.714	2.069	2.398	2.807	3.768
24	0.127	0.390	0.685	1.059	1.318	1.711	2.064	2.391	2.797	3.745
25	0.127	0.390	0.684	1.058	1.316	1.708	2.060	2.385	2.787	3.725
26	0.127	0.390	0.684	1.058	1.315	1.706	2.056	2.379	2.779	3.707
27	0.127	0.389	0.684	1.057	1.314	1.703	2.052	2.373	2.771	3.690
28	0.127	0.389	0.683	1.056	1.313	1.701	2.048	2.368	2.763	3.674
29	0.127	0.389	0.683	1.055	1.311	1.699	2.045	2.364	2.756	3.659
30	0.127	0.389	0.683	1.055	1.310	1.697	2.042	2.360	2.750	3.646
40	0.126	0.388	0.681	1.050	1.303	1.684	2.021	2.329	2.704	3.551
50	0.126	0.388	0.679	1.047	1.299	1.676	2.009	2.311	2.678	3.496
∞	0.126	0.385	0.674	1.036	1.282	1.645	1.960	2.241	2.576	3.291

Table de la distribution Khi-Carrée

Les données dans le tableau sont $\chi^2_\nu(\alpha)$ tel que $P[X > \chi^2_\nu(\alpha)] = \alpha$ si $X \sim \chi^2_\nu$.
colonne = α ; rangée = ν

ν	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Annexe B

La loi normale multivariée

En dimension $p = 1$, on considère d'abord le cas d'une variable aléatoire Z distribuée selon une loi normale centrée réduite $Z \sim \mathcal{N}(0, 1)$. La densité de Z s'écrit alors comme

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)$$

Une variable X de distribution normale avec moyenne μ et variance σ^2 ($X \sim \mathcal{N}(\mu, \sigma^2)$) s'obtient comme $X = \sigma Z + \mu$. La densité de cette variable X s'exprime comme

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}.$$

On considère maintenant la généralisation pour $p \geq 1$ en supposant (Z_1, Z_2, \dots, Z_p) des variables aléatoires indépendantes et identiquement distribuées suivant une loi normale centrée réduite $\mathcal{N}(0, 1)$. On a alors que

$$\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{pmatrix} \sim N_p(\mathbf{0}, \mathbf{I}_p)$$

La densité conjointe des $Z_i, i = 1, \dots, p$, peut alors s'écrire comme

$$\begin{aligned} f(Z_1, \dots, Z_p) &= \prod_{i=1}^p f(z_i) \\ &= \left(\frac{1}{\sqrt{2\pi}} \right)^p \exp \left(- \sum_{i=1}^p z_i^2 / 2 \right) \\ &= \left(\frac{1}{\sqrt{2\pi}} \right)^p \exp \left(- \frac{1}{2} \mathbf{Z}' \mathbf{Z} \right) \end{aligned} \tag{B.1}$$

(B.2)

où p désigne le nombre de variables aléatoires, $\mathbf{0}$ est un vecteur de zéros et \mathbb{I}_p note la matrice identité de dimension $p \times p$. En considérant la transformation de variables $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{(1/2)} \mathbf{Z}$, on obtient que

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

La densité conjointe des X_i est alors donnée par

$$\begin{aligned} f(X_1, \dots, X_p) &= \prod_{i=1}^p f(x_i) \\ &= \left(\frac{1}{\sqrt{2\pi}} \right)^p \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ - \frac{1}{2} (\mathbf{X} - \boldsymbol{\mu})' (\boldsymbol{\Sigma}^{-1/2})' \boldsymbol{\Sigma}^{-1/2} (\mathbf{X} - \boldsymbol{\mu}) \right\} \\ &= \left(\frac{1}{\sqrt{2\pi}} \right)^p \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ - \frac{1}{2} (\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \right\} \end{aligned} \tag{B.3}$$

où

$\boldsymbol{\mu}$ désigne le vecteur contenant la moyenne des p variables,

$\boldsymbol{\Sigma}$ dénote la matrice des variances-covariances de \mathbf{X} (de dimension $p \times p$) (voir la section B.2),

p est le nombre de variables.

On a dans ce cas que le vecteur des variables aléatoires centrées réduites s'obtient comme $\mathbf{Z} = \boldsymbol{\Sigma}^{(-1/2)} (\mathbf{X} - \boldsymbol{\mu})$.

B.1 Espérance et variance

On détermine ci-dessous l'espérance et la variance de $\mathbf{X} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

$$\begin{aligned}\mathbb{E}[\mathbf{X}] &= \mathbb{E}[\boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2} \mathbf{Z}] \\ &= \mathbb{E}[\boldsymbol{\mu}] + \boldsymbol{\Sigma}^{1/2} \mathbb{E}[\mathbf{Z}] \\ &= \boldsymbol{\mu}\end{aligned}$$

$$\begin{aligned}\mathbb{V}\text{ar}[\mathbf{X}] &= \mathbb{V}\text{ar}[\boldsymbol{\Sigma}^{1/2} \mathbf{Z}] \\ &= (\boldsymbol{\Sigma}^{1/2})' \mathbb{V}\text{ar}[\mathbf{Z}] \boldsymbol{\Sigma}^{1/2} \\ &= (\boldsymbol{\Sigma}^{1/2})' \mathbb{I}_p \boldsymbol{\Sigma}^{1/2} \\ &= \boldsymbol{\Sigma}\end{aligned}$$

B.2 La matrice de variances-covariances

La matrice de variances-covariances s'écrit comme

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathbb{V}\text{ar}[X_1] & \mathbb{C}\text{ov}[X_1, X_2] & \cdots & \mathbb{C}\text{ov}[X_1, X_p] \\ \mathbb{C}\text{ov}[X_2, X_1] & \mathbb{V}\text{ar}[X_2] & \cdots & \mathbb{C}\text{ov}[X_2, X_p] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{C}\text{ov}[X_p, X_1] & \mathbb{C}\text{ov}[X_p, X_2] & \cdots & \mathbb{V}\text{ar}[X_p] \end{pmatrix}$$

Dans le cas centré-réduit, cette matrice peut s'écrire comme

$$\boldsymbol{\Sigma} = \mathbb{V}\text{ar}[\mathbf{Z}] = \begin{pmatrix} (e)[2pt, 3cm, 3cm]cccc1 & 0 & \cdots & 0 \\ & 0 & & 1 & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots \\ & 0 & & 0 & \cdots & 1 \end{pmatrix} = \mathbb{I}_p.$$

On note $\sigma_{i,j}$ la valeur correspondant à $\mathbb{C}\text{ov}[X_i, X_j]$. La matrice de variances-covariances possède les deux propriétés suivantes :

1. Elle est symétrique : $\sigma_{i,j} = \sigma_{j,i}$
2. Elle est semi-définie positive :

$$\forall \mathbf{a} \in \mathbb{R}^p, \mathbf{a}' \boldsymbol{\Sigma} \mathbf{a} \geq 0$$

$$\mathbb{V}\text{ar}[\mathbf{a}' \boldsymbol{\Sigma}] \geq 0$$

Elle possède donc p valeurs propres.

Donc si,

- Σ est la matrice de variances-covariances
- Λ est la matrice diagonale des p valeurs propres de Σ
- et que P est la matrice dont les colonnes sont les p vecteurs propres de Σ ,

alors

$$\Sigma = P\Lambda P'$$

et

$$\begin{aligned} |\Sigma| &= |P\Lambda P'| \\ &= |P| |\Lambda| |P'| \\ &= |PP'| |\Lambda| \\ &= \prod_{i=1}^p \lambda_i. \end{aligned}$$

On a donc que $|\Sigma| \neq 0 \Leftrightarrow \lambda_1 > \dots > \lambda_p > 0$ et que Σ^{-1} existe.

Considérons maintenant l'exemple du cas bivarié ($p=2$). La matrice de variances-covariances s'écrit dans ce cas

$$\Sigma = \begin{pmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] \\ \text{Cov}[X_2, X_1] & \text{Var}[X_2] \end{pmatrix}$$

Si $r = \text{corr}(X_1, X_2) = \frac{\text{Cov}[X_1, X_2]}{\sqrt{(\text{Var}[X_1]\text{Var}[X_2])}}$ et $\text{Var}[X_1] = \sigma_1^2$ et $\text{Var}[X_2] = \sigma_2^2$, alors $\text{Cov}[X_1, X_2] = r\sigma_1\sigma_2$.

On a alors

$$\Sigma = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_2\sigma_1 & \sigma_2^2 \end{pmatrix}$$

et son déterminant est

$$|\Sigma| = \sigma_1^2\sigma_2^2 - r^2\sigma_1^2\sigma_2^2.$$

Remarque B.1. On remarque de l'équation précédente que $|\Sigma| \neq 0$ si et seulement si $r \neq \pm 1$. Autrement dit, si $r = \pm 1$, Σ n'est pas inversible, auquel cas la loi normale multivariée (équation B.3) n'a plus de sens. Ce sujet est abordé en détails dans la section sur la multicolinéarité.

Annexe C

Notation matricielle

C.1 Opérations de base sur les matrices

Soit une matrice finie $n \times p$

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & \dots & \dots & a_{2,p} \\ \vdots & \vdots & \dots & \vdots \\ a_{n,1} & \dots & \dots & a_{n,p} \end{pmatrix} = (a_{i,j}).$$

Addition : Soit la matrice $\mathbf{B} = (b_{i,j})$, $n \times p$. Les opérations d'addition $\mathbf{A} + \mathbf{B}$ et $\mathbf{B} + \mathbf{A}$ sont définies :

$$(a_{i,j}) + (b_{i,j}) = (c_{i,j}),$$

où $c_{i,j} = a_{i,j} + b_{i,j}$.

Soustraction : Soit la matrice $\mathbf{B} = (b_{i,j})$, $n \times p$. Les opérations de soustraction $\mathbf{A} - \mathbf{B}$ et $\mathbf{B} - \mathbf{A}$ sont définies :

$$(a_{i,j}) - (b_{i,j}) = (c_{i,j}),$$

où $c_{i,j} = a_{i,j} - b_{i,j}$.

Multiplication : Soit la matrice $\mathbf{D} = (d_{i,j})$, $m \times n$. L'opération $\mathbf{D}\mathbf{A}$ est définie. Le résultat est une matrice $m \times p$, dont l'élément de la ligne i et la colonne j est égal à

$$\sum_{k=1}^n d_{i,k} a_{k,j}.$$

Soit la matrice $\mathbf{E} = (e_{i,j})$, $p \times m$. L'opération $\mathbf{A}\mathbf{E}$ est définie. Le résultat est une matrice $n \times m$, dont l'élément de la ligne i et la colonne j est égal à

$$\sum_{k=1}^m a_{i,k} e_{k,j}.$$

Exemple C.1. On a

$$\begin{pmatrix} 2 & 3 & 0 \\ 7 & 1 & 8 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 4 \\ 9 & 2 & 11 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 3 & 0 \\ 7 & 1 & 8 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 4 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -4 \\ 5 & 0 & 5 \end{pmatrix},$$

et

$$\begin{pmatrix} 2 & 3 & 0 \\ 7 & 2 & 8 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 9 \\ 11 & 23 \end{pmatrix}.$$

□

C.2 Propriétés de base des matrices

Rang : Le rang d'une matrice est le nombre de colonnes (ou de lignes) qui sont linéairement indépendantes.

Matrice Identité : La matrice identité d'ordre p , notée \mathbf{I}_p , est une matrice $p \times p$ composée de 1 sur la diagonale et de zéros ailleurs.

Inverse d'une matrice : Soit une matrice \mathbf{A} , $p \times p$. L'inverse de la matrice \mathbf{A} est notée \mathbf{A}^{-1} et est telle que

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_p.$$

Exemple C.2. Pour une matrice 2×2 , on trouve

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

□

On peut facilement obtenir l'inverse d'une matrice de plus grande dimension à l'aide de la fonction `solve` en R.

Transposition : On note \mathbf{A}^T la transposée de \mathbf{A} . Cette opération consiste en échangeant les lignes et les colonnes, ce qui implique que l'élément de la ligne i et la colonne j de la matrice \mathbf{A}^T est $a_{j,i}$.

Exemple C.3. On a

$$\begin{pmatrix} 2 & 3 & 0 \\ 7 & 1 & 8 \end{pmatrix}^T = \begin{pmatrix} 2 & 7 \\ 3 & 1 \\ 0 & 8 \end{pmatrix}.$$

□

Propriétés des transpositions de matrices :

- $(\mathbf{A}^T)^T = \mathbf{A}$

- $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^\top + \mathbf{B}^\top$
- $(k\mathbf{A})^T = k\mathbf{A}^\top$, si k est un scalaire.
- $(\mathbf{AB})^T = \mathbf{B}^\top \mathbf{A}^\top$

Matrice symétrique : On dit d'une matrice carrée \mathbf{A} qu'elle est symétrique si $\mathbf{A} = \mathbf{A}^T$.

Matrice idempotente : On dit d'une matrice carrée \mathbf{A} qu'elle est idempotente si $\mathbf{A} = \mathbf{AA}$. Si \mathbf{A} est aussi symétrique, on a aussi que $\mathbf{I} - \mathbf{A}$ est symétrique idempotente.

Espérance d'un vecteur aléatoire : Soit $\mathbf{X} = (X_1 \ X_2 \ \dots \ X_n)^T$ un vecteur aléatoire. Alors, l'espérance de \mathbf{X} est

$$E[\mathbf{X}] = (E[X_1] \ E[X_2] \ \dots \ E[X_n])^T.$$

Variance d'un vecteur aléatoire : Soit $\mathbf{X} = (X_1 \ X_2 \ \dots \ X_n)^T$ un vecteur aléatoire. Alors, la variance de \mathbf{X} est

$$Var(\mathbf{X}) = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \dots & Cov(X_2, X_n) \\ \vdots & \vdots & \dots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \dots & Var(X_n) \end{pmatrix}.$$