Exercice 1. 1. Avec la méthode d'Euler, on calcule d'abord y en faisant $L_2 - \frac{1}{\varepsilon} L_1$, on trouve

$$y = \frac{2 - \frac{1}{\varepsilon}}{1 - \frac{1}{\varepsilon}} \ .$$

ce qui est proche de 1 quand ε est petit. Avec une précision de ε, en écrivant $m(x)$ la version "numérique" de x, on a (quand $\varepsilon \to 0$) :

$$m\left(1 - \frac{1}{\varepsilon}\right) = 2 - \frac{1}{\varepsilon} + \delta_1 \left(1 + \frac{1}{\varepsilon}\right) = (1 - \frac{1}{\varepsilon})(1 + \delta_1 + O(\varepsilon))$$

avec $|\delta_1| \leq \varepsilon$. Comme l'inversion préserve l'erreur relative, on a

$$m\left(\frac{1}{1 - \frac{1}{\varepsilon}}\right) = \frac{1}{1 - \frac{1}{\varepsilon}}(1 + \delta_2)$$

où $|\delta_2| = O(\varepsilon)$. En calculant de même l'erreur sur le numérateur, on obtient

$$m(y) = y(1 + \delta_3)$$

où $\delta_3 = O(\varepsilon)$. L'erreur relative sur $m(y)$ reste donc proche de ε. Cependant, la deuxième étape de la méthode d'Euler, en subsituant y dans la première ligne, donne

$$x = \frac{1}{\varepsilon}(1 - y) = \frac{1}{1 - \varepsilon} = 1 + O(\varepsilon) \ .$$

Mais $m(1 - y) = 1 - y + \delta_4(1 + |y|) = 1 - y + O(\varepsilon)$ donc

$$m(x) = 1 + O(1) \ .$$

L'erreur sur x est donc du même ordre que x, ce qui est très mauvais !

2. En intervertissant les deux équations, on obtient

$$y = \frac{1 - 2\varepsilon}{1 - \varepsilon}$$

et

$$x = 2 - y \ .$$

En supposant une précision de ε, on a une erreur relative $O(\varepsilon)$ dans le calcul de y, et donc une erreur absolue en $O(\varepsilon)$ dans le calcul de x. Inverser les lignes permet donc d'avoir une erreur acceptable.

Exercice 2. Soit

$$A = \begin{pmatrix} 1 & 4 \\ -1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} -2 & 3 & 0 \\ 1 & 0 & -4 \\ 2 & 0 & 5 \end{pmatrix} \ .$$

1. On fait une élimination par pivot de Gauss en retenant les opérations. Pour la matrice A, on fait $L_1 \to L_1$ et $L_2 \to L_1 + L_2$, ce qui revient à multiplier par la matrice

$$X_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \ .$$

On a donc

$$X_1 A = \begin{pmatrix} 1 & 4 \\ 0 & 6 \end{pmatrix}$$

donc $A = LU$ avec

$$L = X_1^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \qquad\qquad U = \begin{pmatrix} 1 & 4 \\ 0 & 6 \end{pmatrix}$$

Pour la matrice B, il faut deux opérations :

$$X_1 B = \begin{pmatrix} -2 & 3 & 0 \\ 0 & \frac{3}{2} & -4 \\ 0 & 3 & 5 \end{pmatrix} \qquad X_1 = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad X_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$X_2 X_1 B = \begin{pmatrix} -2 & 3 & 0 \\ 0 & \frac{3}{2} & -4 \\ 0 & 0 & 13 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \qquad X_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$$

d'où on conclut que $B = LU$ avec $U = X_2 X_1 B$ comme ci-dessus et

$$L = X_1^{-1} X_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} .$$

2. On a $\text{dét}(B) = \text{dét}(L)\,\text{dét}(U) = 1 \times (-2) \times \frac{3}{2} \times 8 = -24$.
3. On calcule d'abord $v = Ux$, qui satisfait $Lv = (1, 2, 3)$ ce qui donne $v_1 = 1$, $-\frac{1}{2}v_1 + v_2 = 2$ donc $v_2 = \frac{5}{2}$ et $-v_1 + 2v_2 + v_3 = 3$ donc $v_3 = -1$. Puis on résoud l'équation $Ux = v$, ce qui donne $13u_3 = -1$ (on commence par la dernière ligne), et $\frac{3}{2}u_2 - 4u_3 = \frac{5}{2}$ donc $u_2 = 19/13$, et $-2u_1 + 3u_3 = 1$ donc $u_3 = 22/13$.
4. On a $B^{-1} = U^{-1}L^{-1} = U^{-1}X_2X_1$ donc il suffit d'inverser U, ce qui peut effectivement se faire par la résolution de 3 systèmes linéaires (sinon en terminant le pivot de Gauss).

Exercice 3. Soit α un paramètre réel. On considère la matrice

$$A(\alpha) = \begin{pmatrix} \alpha & 1 \\ 1 & 2 \end{pmatrix}.$$

1. Pour $\alpha \neq 0$ on obtient la factorisation LU de A en une étape,

$$A = \begin{pmatrix} 1 & 0 \\ \frac{1}{\alpha} & 1 \end{pmatrix} \begin{pmatrix} \alpha & 1 \\ 0 & 2 - \frac{1}{\alpha} \end{pmatrix} .$$

Pour $\alpha = 0$ on ne peut avoir une factorisation LU, car $L_{1,1}U_{1,1} = 0$ donc L ou U est de rang < 2, alors que A est de rang 2 pour $\alpha \neq \frac{1}{2}$.

2. Si $\alpha = 0$ Python inverse les lignes, obtenant une décomposition PLU où P est la matrice de permutation correspondant à l'inversion des lignes. Noter que si $|\alpha| < 1$ Python inverse aussi les lignes, afin d'avoir le pivot le plus grand possible (cf exo 1).

Exercice 4. Pour calculer un déterminant naïvement, il faut $n!$ fois n multiplications et $n! - 1$ additions (sans compter le temps de calcul de $\epsilon(\sigma)$ et le temps utilisé pour construire toutes les permutations).

Pour calculer l'inverse par la formule de Cramer, il faut calculer le déterminant de la matrice, puis les n^2 déterminants des mineurs de taille $n-1$, soit $n \times n! + n^2 \times (n-1) \times (n-1)! = n^2 \times n!$ multiplications.
Avec la décomposition LU, à l'utilisation du k-ième pivot on a besoin de $(n-k)^2$ multiplications (ou divisions), soit de l'ordre de $2n^3/6$ multiplications pour le calcul de U. Le calcul de L comme le produit des inverses des matrices de transvection prend également de l'ordre de $2n^3/6$ multiplications.
Pour une matrice $100*100$, le calcul par la méthode de Cramer prend 10^{160} multiplications, donc 10^{150} secondes, alors que la méthode LU demande de l'ordre de 10^6 multiplications, soit 10^{-4} secondes.

Exercice 5. Calculer le déterminant de la matrice de Hilbert de taille 50 puis 100 de manière numérique et de manière exacte. Comparer les résultats et temps de calcul.
Pour quelques matrices M aléatoires de taille 100, 200, 300, calculez le déterminant de M, comparez avec la borne de Hadamard de M (produit des normes des vecteurs colonnes). Qu'observe-t-on ?

Exercice 6. Décomposition de Cholesky.
Soit A une matrice hermitienne définie positive. Il existe une unique matrice triangulaire supérieure C qui a des éléments diagonaux strictement positifs telle que $A = C^*C$ (décomposition de Cholesky).

1. On peut écrire les élements de C comme inconnues puis résoudre les $n(n+1)/2$ équations données par l'égalité $C^*C = A$ successivement.

   Une méthode astucieuse est d'utiliser la décomposition LU. En effet si D est la diagonale de C alors C^*D^{-1} est une matrice triangulaire inférieure avec uniquement des 1 sur la diagonale, et $A = (C^*D^{-1})(DC)$. Ainsi, par unicité de la décomposition LU, on a $L = C^*D^{-1}$ et $U = DC$. Cela suggère l'algorithme suivant :

   (a) calculer U dans la décomposition LU : on trouve $U = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 9 & -3 \\ 0 & 0 & 1 \end{pmatrix}$

   (b) Pour tout i, on divise la i-ième ligne de U par $\sqrt{U_{i,i}}$. On obtient
   $$C = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

   On peut vérifier que la matrice $\tilde{L}$ obtenue en multipliant chaque colonne de L par $\sqrt{U_{i,i}}$ est bien égale à C^*. Ce fait n'est vrai que si A est définie positive. Si A n'est pas symétrique, on a plus $\tilde{L} = C^*$. Si A est symétrique mais pas définie positive, les $U_{i,i}$ ne sont pas tous strictement positifs, et on ne peut donc pas prendre leurs racines. 2) On résout les systèmes $C^*v = b$ puis $Cx = v$ successivement, chacunes de ces résolutions étant rapide puisque C est triangulaire, en résolvant par le bas (en commençant par v_n pour calculer v, puis par le haut (en commençant par x_1) pour calculer x.

2. a) La matrice de l'exo 8 du TD 1 est 2-bande. b) Deux méthodes de résolution : en utilisant l'algorithme (plus naturel) ou par l'absurde (plus astucieux, mais moins intuitif).

   Méthode par construction : Vu l'algorithme décrit ci-dessus, on sait que $C = D^{-\frac{1}{2}}U$ où $A = LU$ est la décomposition LU de A et D est la diagonale de U. Il suffit donc de montrer que U est p-bande. On sait que U est obtenu par un pivot de Gauss descendant.

Pour tout n on note A_n la matrice obtenue dans le pivot de Gauss après n opérations. Montrons par récurrence que A_n est p-bande pour tout n.
On a $A_0 = A$ donc A_0 est p-bande.
De plus, soit $n \geq 0$ tel que A_n soit p-bande. Notons $b_{i,j}$ le coefficient d'indice i, j de A_n et $c_{i,j}$ le coefficient d'indice i, j de A_{n+1}. Pour calculer A_{n+1}, on replace une ligne L_l par $L_l - \lambda L_k$ pour une ligne L_k et un coefficient λ. Le coefficient utilisé dans le pivot est nécessairement $b_{k,k}$, donc $\lambda = \frac{a_{l,k}}{b_{k,k}}$. De plus on fait un pivot descendant, donc $k < l$. Enfin, A_n est p-bande, donc $b_{l,i} = 0$ pour $i \leq l - p$. Ainsi, on peut supposer que $k > l - p$ (sinon $\lambda = 0$ et l'opération est triviale). De plus, puisque $b_{k,k}$ est utilisé comme pivot on a $b_{k,i} = 0$ pour tout $i < k$, donc pour tout $i \leq l - p < k$ on a

$$c_{l,i} = b_{l,i} - \lambda b_{k,i} = 0 \ .$$

Il reste à montrer que $c_{l,i} = 0$ pour tout $i \geq l + p$. Or, pour $i \geq l + p$, on a $b_{l,i} = 0$ et $i \geq k + p$ donc $b_{k,i} = 0$, d'où $c_{l,i} = 0$. Donc $c_{l,i} = 0$ dès que $|l - i| \geq p$, et $c_{j,i} = b_{j,i}$ si $j \neq l$, donc $c_{j,i} = 0$ dès que $|l - i| \geq p$, et donc A_n est p-bande.
Méthode par l'absurde : Supposons que C ne soit pas P-bande, et soit (i, j) avec $|i - j| > p$ tel que $C_{i,j} \neq 0$ avec i minimal (comme C est triangulaire superieur on a alors $i < j$). Alors, en notant $[C^*C]_{i,j}$ le coefficient i, j de C^*C, on a

$$[C^*C]_{i,j} = \sum_{i=1}^{n} \overline{C}_{k,i} C_{k,j}$$

mais C est triangulaire supérieure, donc $C_{k,i} = 0$ pour $k > i$, et comme i est minimal tel que $C_{k,j}$ soit non nul, on a $C_{k,j} = 0$ pour $k < i$. Donc

$$[C^*C]_{i,j} = \overline{C}_{i,i} C_{i,j} \ .$$

Comme A est définie positive, C est inversible, mais $det(C) = \prod_{k=1}^{n} C_{k,k}$ donc $C_{i,i} \neq 0$ et donc $[C^*C]_{i,j} \neq 0$, ce qui est absurde car $A = C^*C$ est p-bande.

Exercice 7. 1. De manière générale, on considère $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ avec $a, b \in \mathbb{R}$. Cette matrice est symétrique, donc elle admet une base orthonormale de vecteurs propres. On montre facilement qu'on peut prendre comme vecteurs propres $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ avec valeurs propres respectives $\lambda_1 = a + b$ et $\lambda_2 = a - b$.

2. Notons Q_1 le projecteur orthogonal sur l'espace propre $\mathbb{R}v_1$ et $Q_2 = I - Q_1$ le projecteur orthogonal sur $\mathbb{R}v_2$. On a alors

$$A = \lambda_1 Q_1 + \lambda_2 Q_2$$

et donc $A^{-1} = \lambda_1^{-1} Q_1 + \lambda_2^{-1} Q_2$, d'où

$$x = \lambda_1^{-1} Q_1 b + \lambda_2 Q_2 b \ .$$

3. En prenant $b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \sqrt{2} v_1$ on a $Q_1 b = b$ et $Q_2 b = 0$, donc $x = \lambda_1^{-1} b = \frac{1}{2001} b$.

Pour obtenir une grande erreur relative sur x, on choisit une erreur sur b égale à $10^{-2}\sqrt{2} v_2$. On a bien une erreur relative sur b petite :

$$\frac{\|\Delta b\|}{\|b\|} = 10^{-2}$$