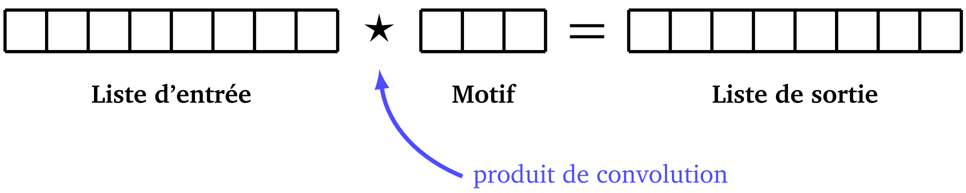
Convolution : une dimension

Vidéo ■ partie 11. Convolution : une dimension

Ce chapitre permet de comprendre la convolution dans le cas le plus simple d'un tableau à une seule dimension.

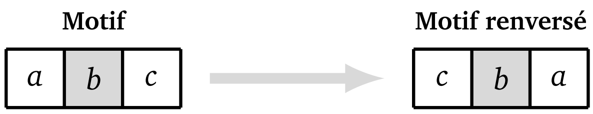
1. Idée

La convolution est une opération qui à partir d'un tableau de nombres et d'un motif produit un nouveau tableau de nombres.

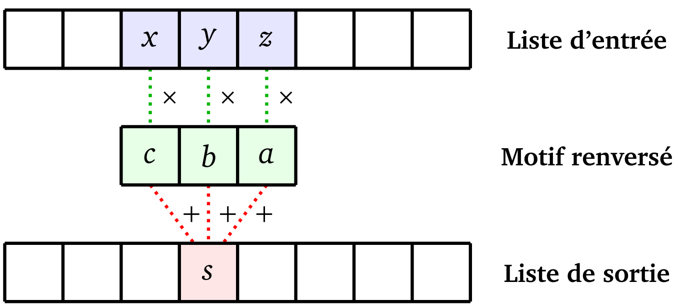


Le calcul de la liste de sortie se fait par des opérations très simples.

- On commence par renverser le motif.

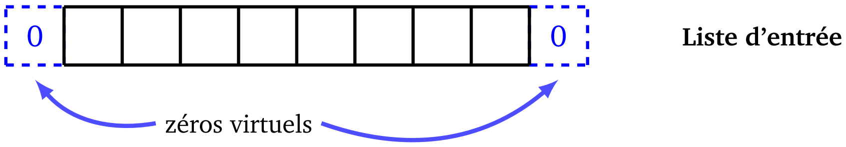


- On calcule la liste de sortie terme par terme :
  — on centre le motif renversé sous la liste d'entrée, à la position à calculer,
  — on multiplie terme à terme les éléments de la liste d'entrée et ceux du motif,
  — la somme de tous ces produits est le terme de la liste de sortie.



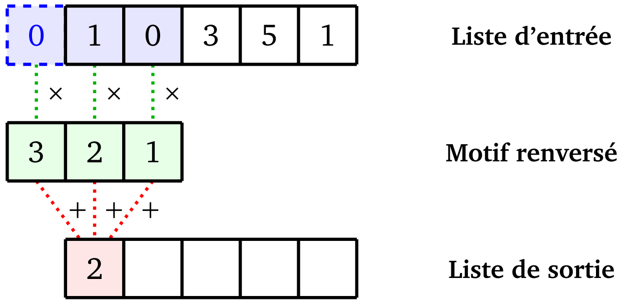
$$s = xc + yb + za$$

Pour le calcul des coefficients sur les bords, on rajoute des zéros virtuels à gauche et à droite de la liste d'entrée.

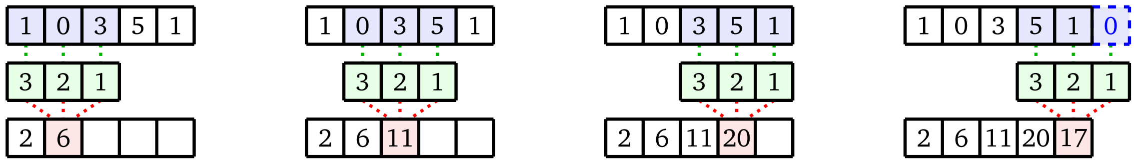


Notation. On note $f \star g$ ce ***produit de convolution tronqué***. La liste de sortie est de même taille que celle de la liste d'entrée.

Exemple.
Calculons la convolution $[1,0,3,5,1] \star [1,2,3]$. On commence par calculer le coefficient le plus à gauche (après ajout d'un zéro virtuel à la liste d'entrée).

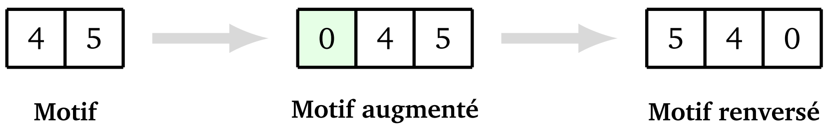


On continue en calculant les coefficients un par un.

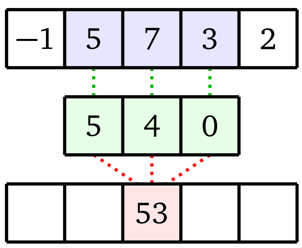


Ainsi $[1,0,3,5,1] \star [1,2,3] = [2,6,11,20,17]$.

Exemple.
Si le motif est de longueur paire alors on ajoute un 0 au début du motif pour le calcul. Ainsi $[-1,5,7,3,2] \star [4,5] = [-1,5,7,3,2] \star [0,4,5]$



Lorsque l'on renverse le motif, le zéro ajouté se retrouve à droite !



Après calcul on obtient $[-1,5,7,3,2] \star [4,5] = [-4,15,53,47,23]$.

2. Exemples

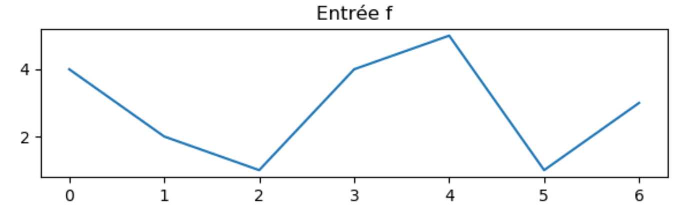
Voyons des exemples de motifs et leur effet sur des tableaux.

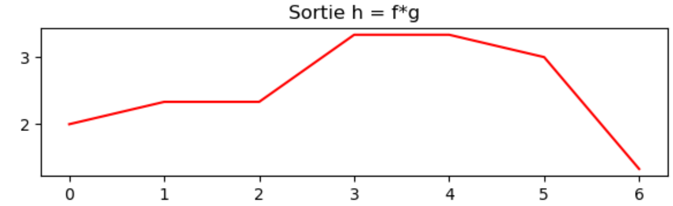
Moyenne mobile. La convolution par le motif $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$ correspond à effectuer une moyenne mobile sur trois termes.

Exemple :

$$[4,2,1,4,5,1,3] \star \frac{1}{3}[1,1,1] \simeq [2.00,2.33,2.33,3.33,3.33,3.00,1.33]$$

On peut représenter graphiquement une liste $[x_0, x_1, \ldots, x_N]$ en reliant les points de coordonnées (k, x_k) pour k entre 0 et N. Voici la représentation du tableau d'entrée en bleu et celle du produit de convolution en rouge.



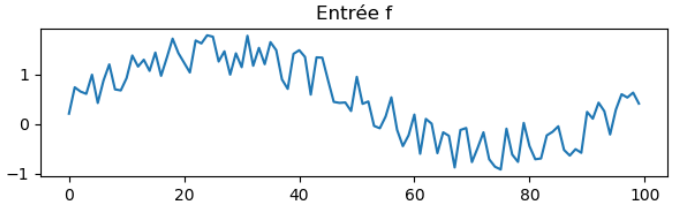


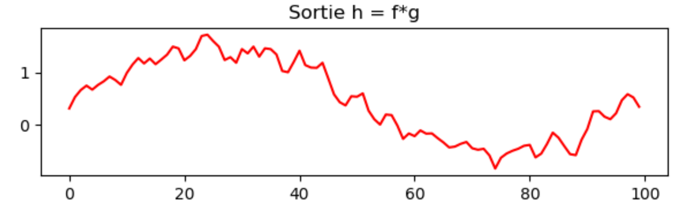
Plus généralement, le motif de longueur n : $\frac{1}{n}[1,1,1,\ldots,1]$ correspond à une moyenne mobile sur n termes. Cela permet de « lisser » des données et de supprimer du « bruit ».

Exemple :

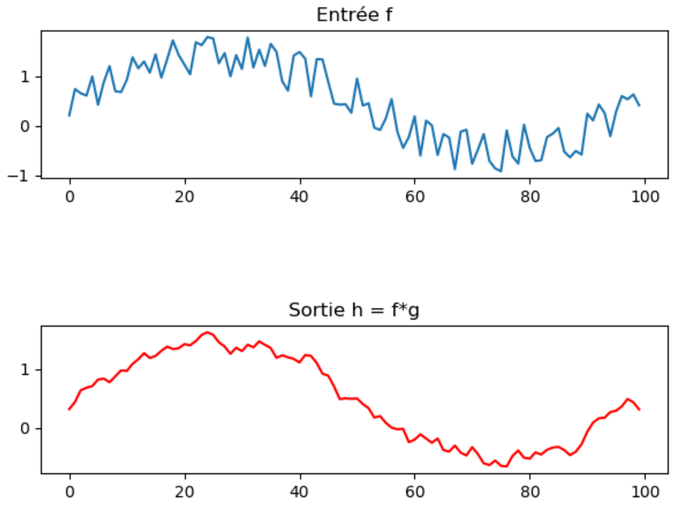
$$[4,2,1,4,5,1,3] \star \frac{1}{5}[1,1,1,1,1] = [1.4,2.2,3.2,2.6,2.8,2.6,1.8]$$

Voici la représentation d'un tableau de 100 points, au comportement assez chaotique, ainsi que le résultat de son produit de convolution par le motif $\frac{1}{3}[1,1,1]$: faire une moyenne mobile « lisse » la courbe.





En changeant maintenant le motif de convolution en $\frac{1}{5}[1,1,1,1,1]$ la courbe est encore plus lisse.



Homothétie. La convolution par le motif $[k]$ (de longueur 1 et qui pourrait encore s'écrire $[0,k,0]$) multiplie tous les termes de la liste d'entrée par un facteur k.
Exemple :

$$[1,2,3,4,5,6,7,8,9] \star [2] = [2,4,6,8,10,12,14,16,18].$$

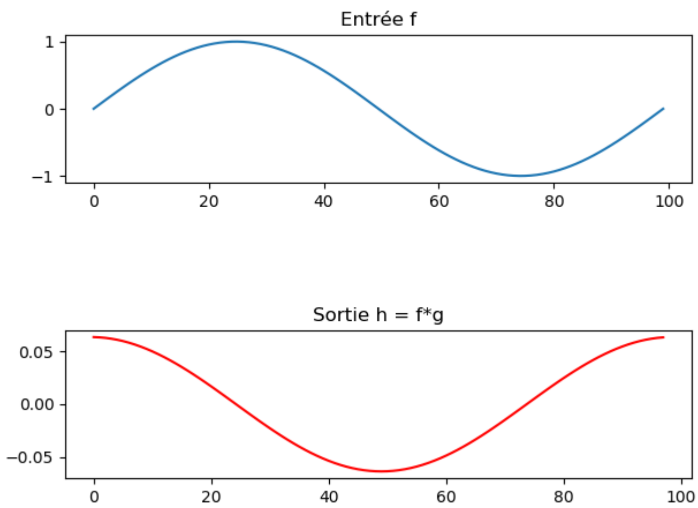
Translation. La convolution par le motif $[0,0,1]$ décale la liste d'un cran vers la droite (avec ajout d'un zéro en début, le dernier terme disparaissant).
Exemple :

$$[1,2,3,4,5,6,7,8,9] \star [0,0,1] = [0,1,2,3,4,5,6,7,8].$$

Dérivée. La convolution par le motif $[0,1,-1]$ (qui s'écrit aussi $[1,-1]$) correspond au calcul de la dérivée de la liste d'entrée vu comme une fonction $n \mapsto f(n)$.
Exemple :

$$[16,9,4,1,0,1,4,9,16] \star [0,1,-1] = [16,-7,-5,-3,-1,1,3,5,7].$$

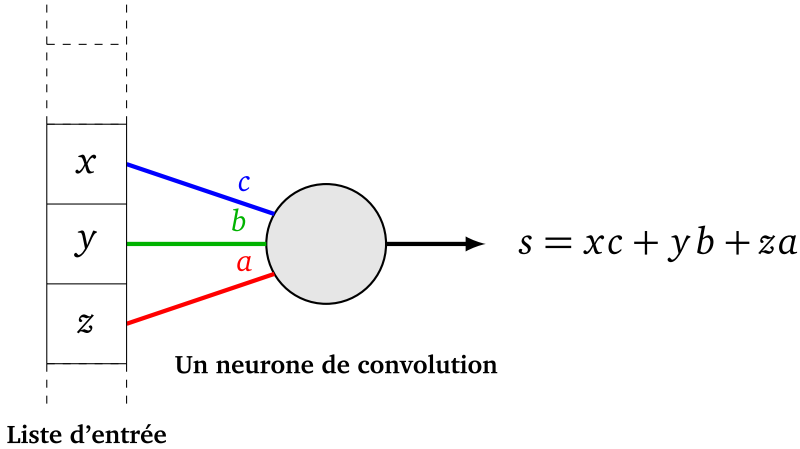
Le calcul correspond à $f(n+1)-f(n)$, une analogie discrète du taux d'accroissement $\frac{f(x+h)-f(x)}{h}$ avec $h=1$. Pour un tableau d'entrée de 100 valeurs $[\sin(0), \sin(2\pi/100), \sin(4\pi/100), \ldots, \sin(198\pi/100)]$, définies par la fonction sinus sur $[0,2\pi]$, le produit de convolution par le motif $[0,1,-1]$ donne logiquement un graphe ressemblant au cosinus (à un facteur multiplicatif près).



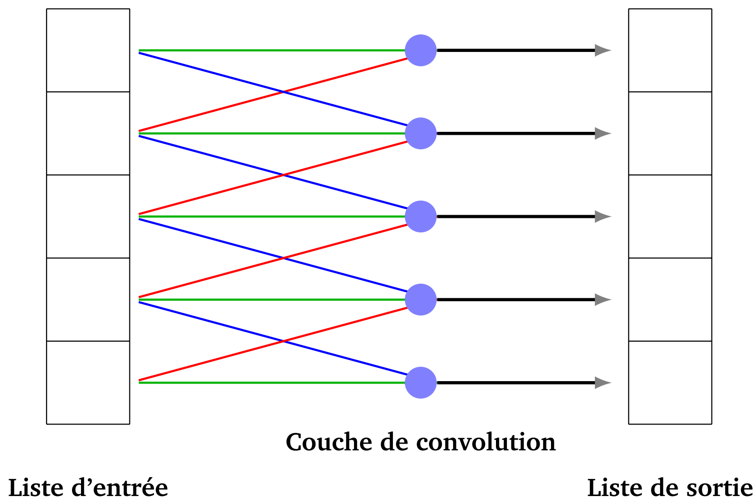
On retient donc que la convolution par un motif, peut rendre la liste plus lisse, la transformer (homothétie, translation) et peut en extraire certaines propriétés (comme la dérivée).

3. Réseau de neurones

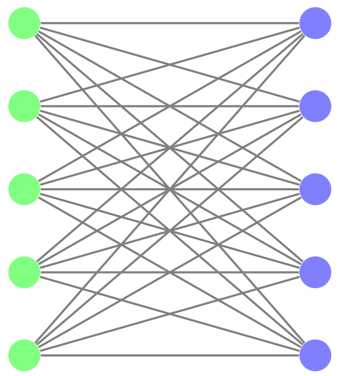
Comment créer un nouveau type de réseau de neurones réalisant un produit de convolution ? Nous allons créer un nouveau type de neurone et un nouveau type de couche de neurones correspondant à un motif donné.
Pour un motif $[a, b, c]$, un neurone de convolution possède trois arêtes, de poids a, b, c, et est relié à seulement trois valeurs de l'entrée.



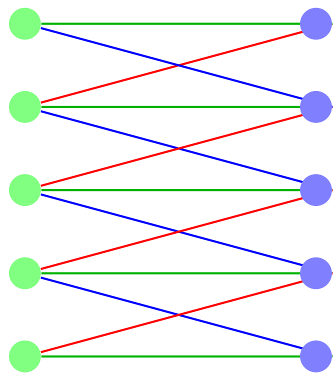
La sortie est $s = xc + yb + za$. On pourrait aussi composer par une fonction d'activation ϕ, auquel cas la sortie serait $s = \phi(xc + yb + za)$.
On peut représenter une couche de convolution par un ensemble de neurones de convolution, ceux-ci ayant tous les mêmes poids $[a, b, c]$ et chacun étant relié à seulement trois valeurs de l'entrée. La sortie produite est alors le produit convolution de l'entrée par le motif $[a, b, c]$. Pour une entrée de longueur n, la couche de convolution possède n neurones, chaque neurone ayant trois arêtes (sauf les neurones de bord). Mais au total, il n'y a que trois poids différents.



Quel est l'avantage de cette couche de neurones par rapport à une couche complètement connectée (comme dans les chapitres précédents) ? Dans une couche complètement connectée de n neurones ayant une entrée de taille n, il y a n^2 coefficients alors qu'ici il n'y a que 3 coefficients (quel que soit n). Par exemple, si $n = 100$ alors il y a seulement 3 coefficients à calculer au lieu de 10 000.



Couche complètement connectée



Couche de convolution

Sur la figure du gauche, nous avons une couche complètement connectée à la précédente. Chaque arête correspond à un poids à calculer. Sur la figure de droite, une couche de convolution (de taille 3) est connectée à la précédente. Chaque neurone n'a que 3 arêtes (sauf les neurones extrêmes qui n'en ont que deux) et de plus les arêtes partagent leur poids (les arêtes rouges ont toutes le même poids a, les arêtes vertes le même poids b, les arêtes bleues le même poids c).

4. Définition mathématique

Cette partie n'est pas utile pour les chapitres suivants, mais présente les choses de façon plus théorique. En particulier, il n'y a plus à ajouter de zéros virtuels (ni sur la liste d'entrée, ni sur un motif de longueur paire).

Définition.
Soient $(f(n))_{n\in\mathbb{Z}}$ et $(g(n))_{n\in\mathbb{Z}}$ deux suites de nombres réels. Le *produit de convolution* $f \,\widetilde{\star}\, g$ est la suite $(h(n))_{n\in\mathbb{Z}}$ dont le terme général est défini par :

$$\boxed{f \,\widetilde{\star}\, g(n) = \sum_{k=-\infty}^{+\infty} f(n-k)\cdot g(k)}$$

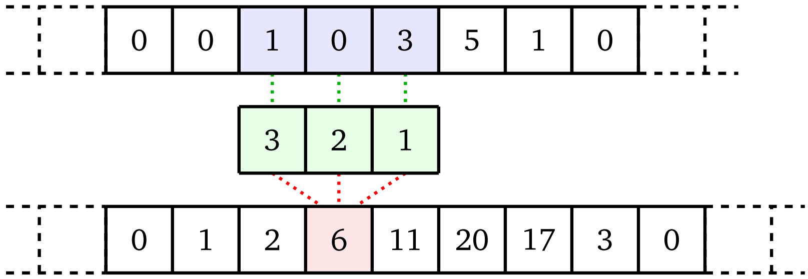
Il ne faut pas avoir peur de cette formule. En particulier, dans les situations rencontrées ici, il n'y a pas vraiment une infinité de termes à calculer. Voici une formule plus simple, lorsque l'on suppose que les termes de g sont nuls en dehors des indices appartenant à $[-K,+K]$:

$$\boxed{f \,\widetilde{\star}\, g(n) = \sum_{k=-K}^{+K} f(n-k)\cdot g(k)}$$

Reprenons l'exemple de la convolution de $[1,0,3,5,1]$ par $[1,2,3]$. Choisissons $(f(n))_{n\in\mathbb{Z}}$ la suite dont les termes non tous nuls sont $[f(0)=1, f(1)=0, f(2)=3, f(3)=5, f(4)=1]$ et choisissons $(g(n))_{n\in\mathbb{Z}}$ la suite dont les termes non tous nuls sont $[g(-1)=1, g(0)=2, g(1)=3]$. Alors on obtient la convolution $h = f \,\widetilde{\star}\, g$ par la formule ci-dessus :

$$f \,\widetilde{\star}\, g(n) = \sum_{k=-1}^{+1} f(n-k)\cdot g(k) = f(n-1)g(1) + f(n)g(0) + f(n+1)g(-1).$$

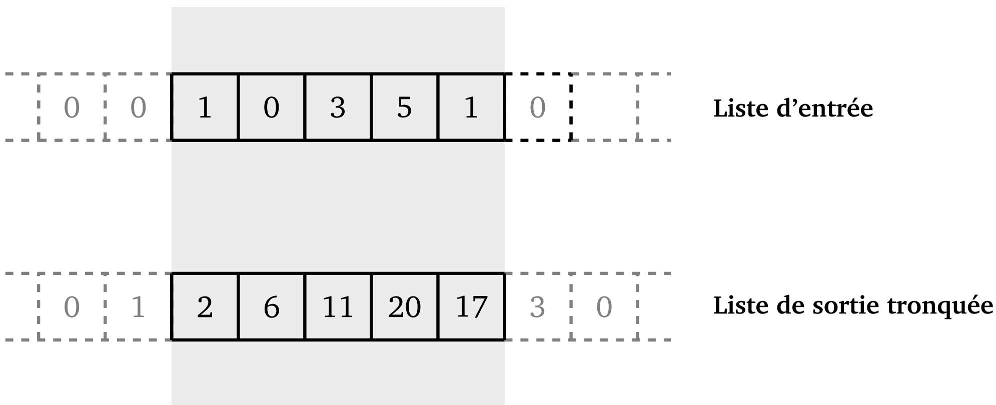
On retrouve la méthode pratique vu précédemment : on renverse le motif g avant de faire une série de multiplications. Par exemple $f \,\widetilde{\star}\, g(1) = f(0)g(1) + f(1)g(0) + f(2)g(-1) = 1\times 3 + 0\times 2 + 3\times 1 = 6$.



Le résultat de la convolution est la suite

$$[\ldots, 0, 0, 1, 2, 6, 11, 20, 17, 3, 0, 0, \ldots]$$

où le terme de rang 0 est $f \tilde{\star} g(0) = 2$. Ce résultat est indépendant, à un décalage près, des choix opérés dans les définitions de f et g.
La convolution tronquée consiste à ne retenir qu'une sous-liste de même taille que la liste d'entrée en tronquant les bords (la façon de tronquer dépend du choix des définitions de f et g). Ainsi : $[1, 0, 3, 5, 1] \star [1, 2, 3] = [2, 6, 11, 20, 17]$.



Terminons par une propriété remarquable du produit de convolution. En réindexant la somme de la définition, on obtient :

$$f \tilde{\star} g(n) = \sum_{k=-\infty}^{+\infty} f(k) \cdot g(n-k).$$

Ce qui signifie que le produit de convolution est symétrique :

$$\boxed{f \tilde{\star} g(n) = g \tilde{\star} f(n)}$$

Remarque.

- La convolution est très utilisée en théorie du signal, c'est pourquoi on trouve aussi le vocabulaire *signal d'entrée*, *signal de sortie*. C'est aussi pour des raisons de signal qu'on renverse une des listes avant de faire les multiplications.
- Le ***motif*** s'appelle aussi ***noyau*** (*kernel* en anglais) mais aussi ***masque*** ou ***filtre***.