# Approximation et Interpolation Polynomiale, Application à l'intégration numérique.

## Laurent RAYMOND

## 14 février 2006

## Table des matières

1	Motivations						
<b>2</b>	Rég	ression Polynomiale	2				
	2.1	Régression linéaire	2				
	2.2	Régression polynômiale	5				
	2.3	Exemples	5				
		2.3.1 Regression par une parabole	6				
		2.3.2 Régression linéaire	6				
3	Inte	Interpolation polynomiale.					
	3.1	Méthode de Lagrange	8				
	3.2	Exemples	9				
		3.2.1 Echantillonage régulier	9				
		3.2.2 Points de Gauss	9				

#### 1 Motivations

Les méthodes d'approximation dont nous allons traiter dans ce chapitre sont souvent utilisées par exemple en science expérimentale ou en statistique. En effet, le calcul de certaines fonctions de manière précise peut parfois présenter de réelles difficultés et demander des temps très long, alors que l'on peut se contenter de valeurs approchées. On peut dans ce cas avoir recours à une approximation de ces objets compliquées par des fonctions simples. Un exemple de cette stratégie sera exposé dans un chapitre de ce cours, traitant de la transformée de Fourier discrète. Dans ce chapitre, les fonctions simples utilisées comme base seront des polynômes, dans le cas de la décomposition en série de Fourier, on utilise des polynômes trigonométriques.

Par ailleurs, il existe de très nombreux cas, où on ne connait une fonction que par ses valeurs en un certain nombre de points. Par exemple, lorsqu'on effectue une série de mesures, on obtient les valeurs de la grandeur mesurée Y en fonction du paramètre expérimental X que l'on fait varier. On n'a en règle générale pas accès à la fonction F qui relie ce paramètre à la valeur mesurée, mais il peut être intéressant de pouvoir prédire une valeur approchée de F(x) pour des valeurs de x que l'on n'a pas mesuré (si la valeur de x se trouve entre 2 valeurs mesurées, on parle d'interpolation, si par contre, la valeur se trouve a l'extérieur de toutes les mesures, on parle d'extrapolation). Un autre exemple est celui d'une fonction que l'on connait parfaitement, mais dont le calcul prend un temps très long (par exemple, dans les calculatrices, les fonctions trigonométriques sont probablement approchées par des expressions polynomiales).

Une utilisation courante d'approximation par un polynôme est la régression linéaire. Nous en rappelons brièvement le principe dans la partie 2. Un autre exemple est l'interpolation de la partie 3 qui permet d'obtenir une courbe passant par un ensemble de points déterminés à l'avance. Nous avons choisi d'illustrer l'utilisation de ces méthodes dans la partie ?? traitant du calcul numérique d'intégrale.

## 2 Régression Polynomiale

On imagine une expérience dépendant d'un paramètre X et conduisant à une valeur mesurée Y. Par exemple, si on mesure l'intensité (Y) traversant une résistance en fonction de la tension (X) à ses bornes.

A la fin de cette expérience, on obtient un ensemble de valeurs du paramètre  $(X_i)_{i=1..N}$ , et de la valeur de la mesure associée  $(Y_i)_{i=1..N}$ . On suppose de plus que la loi qui relie Y à X est une fonction polynomiale. On sait bien que lorsqu'on mesure une grandeur, les valeurs obtenues sont généralement approchées, et que donc, si on représente les couples  $(X_i, Y_i)$ , par des points du plans, ceux-ci se dispersent (plus ou moins selon les conditions expérimentales et la qualité de l'expérimentateur) autour de la courbe y = f(x). Dans le cas de l'exemple de la résistance, on obtiendra un nuage de point qui devrait ne pas être très éloigné d'une droite passant par l'origine dont la pente est la valeur de la résistance.

## 2.1 Régression linéaire.

On suppose qu'on peut approcher le nuage de points obtenu lors de l'expérience de manière satisfaisante par une relation entre x et y de la forme y = ax + b où a et b sont des paramètres à déterminer. Pour choisir "la meilleure" approximation, on doit se donner un

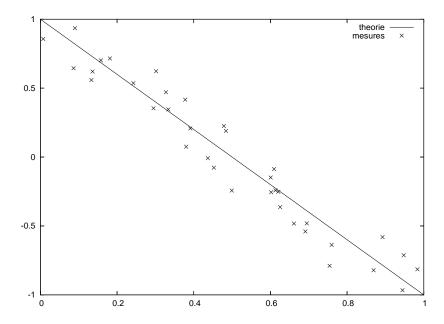


FIG. 1 – Dispersion des points expérimentaux autour de la valeur théorique. Ici, la loi est y = -2x + 1, et il y a des erreurs de mesure sur x et sur y.

critère de qualité. L'ingrédient de base est la définition d'une distance entre le nuage de points et la droite. On cherchera à la rendre minimale en ajustant les paramètres a et b.

On rappelle que la distance entre un point M(x,y) et une droite est en général défini par le minimum de l'ensemble de toutes les distances entre M et les points de la droite. Si la distance entre 2 points est la distance Euclidienne usuelle dans le plan, ce minimum est atteint lorsque l'on choisit la projection orthogonale du point M sur la droite. Pour les besoins de la régression linéaire, en fonction des conditions expérimentales, on adopte en général un choix légèrement différent de distance.

#### On peut utiliser:

- Mesure des projections sur l'axe des y: lorsqu'on considère que le paramètre expérimental x est connu avec une grande précision, on peut supposer que seule la mesure y est incertaine. On mesure alors la distance entre le point de coordonnées  $(x_i, y_i)$  et la droite d'équation y = ax + b par le nombre  $e_i$  tel que  $e_i^2 = (ax_i + b y_i)^2$ .
- Mesure des projections sur l'axe x: ici par contre, on considère que la mesure effectuée est fiable, alors que les conditions de l'expérience sont plus incertaines, on mesure donc la distance par  $f_i^2 = [(y_i b)/a x_i]^2$ . Notons que ceci correspond à l'échange entre x et y, avec l'utilisation des paramètres  $\tilde{a} = 1/a$  et  $\tilde{b} = -b/a$ .
- Mesure de la distance usuelle : on accorde autant de confiance aux valeurs de  $x_i$  et de  $y_i$ , la mesure de distance est obtenue par  $d_i^2 = e_i^2 + f_i^2$ .

Dans les 3 cas précédents, il est équivalent de rechercher le minimum de la distance ou le minimum de son carré. Techniquement, il est préférable d'utiliser des fonctions dérivables lorsqu'on cherche un extrémum, il est donc plus simple d'utiliser les carrés des distances.

Dans le cas d'un ensemble de N points  $(x_i, y_i)_{i=0..N-1}$ , on définit :

$$S(a,b) = \sum_{i=0}^{N-1} e_i^2 = \sum_{i=0}^{N-1} (ax_i + b - y_i)^2,$$
 (1)

$$T(a,b) = \sum_{i=0}^{N-1} f_i^2 = \sum_{i=0}^{N-1} [(y_i - b)/a - x_i]^2,$$
 (2)

$$U(a,b) = S(a,b) + T(a,b).$$
 (3)

Les 3 versions du problème de régression linéaire correspondent donc à la recherche d'un minimum de l'une de ces grandeurs (S, T ou U) lorsqu'on fait varier les paramètres a et b. Dans ce qui suit, nous ne considèrerons que le premier cas de régression, qui revient à rechercher un minimum de S. Pour le trouver, on peut tenter d'annuler ses dérivées partielles :

$$0 = \frac{\partial S}{\partial a}(a,b) = \sum_{i=0}^{N-1} 2x_i(ax_i + b - y_i) = 2\left(a\sum_{i=0}^{N-1} x_i^2 + b\sum_{i=0}^{N-1} x_i - \sum_{i=0}^{N-1} x_i y_i\right), \quad (4)$$

$$0 = \frac{\partial S}{\partial b}(a,b) = \sum_{i=0}^{N-1} 2(ax_i + b - y_i) = 2\left(a\sum_{i=0}^{N-1} x_i + b\sum_{i=0}^{N-1} 1 - \sum_{i=0}^{N-1} y_i\right).$$
 (5)

Posons

$$\langle x \rangle = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad \langle y \rangle = \frac{1}{N} \sum_{i=0}^{N-1} y_i$$
 $\langle x^2 \rangle = \frac{1}{N} \sum_{i=0}^{N-1} x_i^2 \quad \langle y^2 \rangle = \frac{1}{N} \sum_{i=0}^{N-1} y_i^2 \quad \langle xy \rangle = \frac{1}{N} \sum_{i=0}^{N-1} x_i y_i$ 

En divisant les equations (4) et (5) par 2N et en remplaçant par ces dernières expressions, on obtient les valeurs usuelles des coefficients de régression linéaire :

$$a = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2} \quad b = \langle y \rangle - a \langle x \rangle . \tag{6}$$

On est bien entendu en droit de se demander pourquoi la valeur de S est alors minimale. En fait il suffit de montrer que la valeur de S ne peut qu'augmenter lorsqu'on varie a et b autour de leurs valeurs obtenues à l'équation (6).

Le cas de b se résoud simplement : Supposons que l'on ait une valeur de a fixée. Si b diminue suffisamment, tous les points de mesures se trouvent au dessus de la droite et une diminution supplémentaire de b implique une augmentation de la distance entre la droite et le nuage de points, puisque toutes les distances augmentent. De même si on augmente beaucoup b, on augmente nécessairement S à partir du moment où tous les points se trouvent sous la droite. On peut donc conclure que S tend vers  $+\infty$  lorsque b tend vers  $\pm \infty$ , il existe donc **au moins** un **minimum** entre les deux. Or la dérivée partielle de S par rapport à b ne s'annule que pour  $b = \langle y \rangle - a \langle x \rangle$ . Il existe donc **au plus une valeur extrémale** de S en fonction de b, qui est donc le minimum de S. Il reste à montrer que pour cette valeur de b, un extrémum de  $S(a, \langle y \rangle - a \langle x \rangle)$  est en fait un minimum. Pour celà, il suffit de calculer les 2 premières dérivées de la fonction  $a \mapsto S(a, \langle y \rangle - a \langle x \rangle)$ , pour vérifier que la dérivée seconde est positive lorsque la dérivée première s'annule. Ce dernier calcul est laissé en exercice.

#### 2.2 Régression polynômiale.

Cette procédure de régression peut se généraliser en utilisant un polynôme P de degré quelconque p plutôt qu'une fonction affine. Bien que plus lourds, les calculs sont aussi directs que précedemment. On écrit simplement l'annulation des dérivées partielles de la fonction S. Le nombre de paramètres à ajuster est maintenant égal au nombre de coefficients du polynôme P soit un de plus que son degré p.

On a donc à déterminer l'ensemble des  $(a_k)_{k=0..p}$  qui définissent le polynôme P par

$$P(X) = a_0 + a_1 X + \ldots + a_p X^p = \sum_{k=0}^p a_k X^k.$$

Les valeurs des coefficients  $a_k$  doivent minimiser la distance S définie par

$$S(a_0, \dots, a_p) = \sum_{i=0}^{N-1} (P(x_i) - y_i)^2 = \sum_{i=0}^{N-1} \left[ \left( \sum_{k=0}^p a_k x_i^k \right) - y_i \right]^2.$$
 (7)

Un calcul similaire à celui de la section (2.1) conduit à :

$$\frac{\partial S}{\partial a_j}(a_0, \dots, a_p) = \sum_{i=0}^{N-1} 2x_i^j \left[ \left( \sum_{k=0}^p a_k x_i^k \right) - y_i \right] = 2 \left[ \sum_{k=0}^p a_k \sum_{i=0}^{N-1} x_i^j x_i^k - \sum_{i=0}^{N-1} x_i^j y_i \right]. \tag{8}$$

On définit la matrice Q à N lignes et p+1 colonnes par  $Q_{ij} = Q_{ji}^* = x_{i-1}^{j-1}$ , i=1..N, j=1..(p+1). On notera que W le produit  $Q^*Q$  est bien défini et que c'est une matrice carrée symétrique de taile p+1 telle que

$$W_{kk'} = (Q^*Q)_{kk'} = \sum_{i=1}^{N} Q_{ki}^* Q_{ik'} = \sum_{i=0}^{N-1} x_i^{(k-1)+(k'-1)},$$

avec k = 1..(p+1) et k' = 1..(p+1). On peut remarquer que

$$\begin{pmatrix} \frac{\partial S}{\partial a_0}(a_0, \dots, a_p) \\ \vdots \\ \frac{\partial S}{\partial a_p}(a_0, \dots, a_p) \end{pmatrix} = 2Q^*Q \begin{pmatrix} a_0 \\ \vdots \\ a_p \end{pmatrix} - 2Q^* \begin{pmatrix} y_0 \\ \vdots \\ y_{N-1} \end{pmatrix}.$$

Pour annuler ces dérivées partielles, il reste donc à résoudre le système d'équations linéaires

$$Q^*QA = Q^*Y,$$

où A est le vecteur colonne défini par  $A_k = a_{k-1}$  pour k = 1..(p+1) et Y vecteur colonne défini par  $Y_i = y_{i-1}$  pour i = 1..N. Si on définit  $B = Q^*Y$ , l'équation prend la forme simple d'un système de p+1 équations à p+1 inconnues :

$$WA = B \tag{9}$$

#### 2.3 Exemples.

Pour illustrer un peu cet exposé aride, nous allons utiliser le logiciel de calcul octave ou matlab pour réaliser ce calcul dans un cas concret.

#### 2.3.1 Regression par une parabole.

Nous allons calculer l'équation d'une parabole passant "au plus près" d'un ensemble de données expérimentales :

i	1	2	3	4	5
					0.500
$\overline{y_i}$	0.225	0.209	0.200	0.221	0.259

D'après la section précédente, il suffit de calculer la matrice W et le vecteur B:

$$W = \begin{pmatrix} \sum_{i=1}^{5} x_i^0 & \sum_{i=1}^{5} x_i^1 & \sum_{i=1}^{5} x_i^2 \\ \sum_{i=1}^{5} x_i^1 & \sum_{i=1}^{5} x_i^2 & \sum_{i=1}^{5} x_i^3 \\ \sum_{i=1}^{5} x_i^2 & \sum_{i=1}^{5} x_i^3 & \sum_{i=1}^{5} x_i^4 \end{pmatrix} = \begin{pmatrix} 5.0000 & 1.5000 & 0.5500 \\ 1.5000 & 0.5500 & 0.2250 \\ 0.5500 & 0.2250 & 0.0979 \end{pmatrix}$$

$$B = \begin{pmatrix} \sum_{i=1}^{5} x_i^0 y_i \\ \sum_{i=1}^{5} x_i^1 y_i \\ \sum_{i=1}^{5} x_i^2 y_i \end{pmatrix} = \begin{pmatrix} 1.11400 \\ 0.34220 \\ 0.12872 \end{pmatrix}$$

On en déduit en utilisant une méthode de résolution au choix que

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.26780 \\ -0.51143 \\ 0.98571 \end{pmatrix}$$

Que l'on peut représenter par la courbe suivante :

#### 2.3.2 Régression linéaire.

Il est intéressant de vérifier que l'on retrouve bien la formule de la régression linéaire que l'on a établie précédemment, en utilisant ce formalisme plus général.

Pour celà, il suffit de calculer les matrices W et B dans le cas particulier p=1, on a alors :

$$W = \begin{pmatrix} \sum_{i=1}^{N} x_i^0 & \sum_{i=1}^{N} x_i^1 \\ \sum_{i=1}^{N} x_i^1 & \sum_{i=1}^{N} x_i^2 \end{pmatrix} = \begin{pmatrix} N & N < x > \\ N < x > & N < x^2 > \end{pmatrix}$$

et

$$B = \begin{pmatrix} \sum_{i=1}^{N} x_i^0 y_i \\ \sum_{i=1}^{N} x_i^1 y_i \end{pmatrix} = \begin{pmatrix} N < y > \\ N < xy > \end{pmatrix}$$

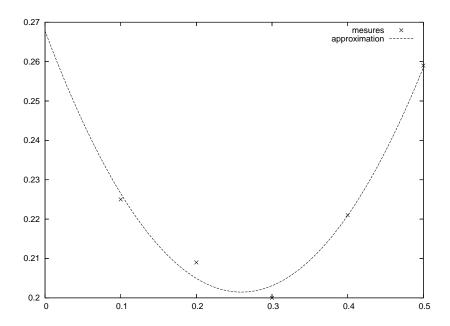


FIG. 2 – Régression par un polynôme de degré 2. L'incertitude est supposée uniquement sur les valeurs y.

On vérifie donc facilement que le vecteur

$$A = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{N^2(\langle x^2 > -\langle x >^2)} \begin{pmatrix} N < x^2 > & -N < x > \\ -N < x > & N \end{pmatrix} \begin{pmatrix} N < y > \\ N < xy > \end{pmatrix}$$

et donc

$$\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \langle y \rangle - \langle x \rangle & \langle xy \rangle - \langle x \rangle \langle y \rangle \\ & \langle xy \rangle - \langle x \rangle \langle y \rangle \\ & \langle x^2 \rangle - \langle x \rangle \langle y \rangle \end{pmatrix}$$

On rappelle qu'une matrice carrée C de taille 2 de coefficients  $c_{ij}$  est inversible si et seulement si son déterminant  $Det(C) = c_{11}c_{22} - c_{12}c_{21}$  est non nul. Son inverse est obtenu explicitement par

$$C^{-1} = \frac{1}{Det(C)} \begin{pmatrix} c_{22} & -c_{12} \\ -c_{12} & c_{11} \end{pmatrix}$$

## 3 Interpolation polynomiale.

Le cas où le nombre de coefficients du polynôme utilisé pour l'approximation et égal au nombres des points de mesure est un autre cas particulier intéressant du calcul précédent qui mérite qu'on s'y attarde. Ce dernier prend alors le nom d'interpolation polynomiale car non seulement le polynôme est la meilleure approximation possible, mais de plus la distance entre le nuage de points et la courbe est alors nulle. Autrement dit, la courbe obtenue passe par tous les points de mesure.

Si on reprend le calcul de la section 2.2 avec p = N, on obtient uniquement des matrices carrées, et le calcul est alors réduit à la résolution du système d'équations suivants :

$$Q^*QA = Q^*Y \tag{10}$$

où la matrice Q est la matrice carrée dite de Van Der Monde dont les coefficients sont obtenus par  $Q_{i,j} = x_{i-1}^{j-1}$ . On peut montrer par récurrence que le déterminant de cette matrice est proportionnel au produit des facteurs  $(x_i - x_{i'})$  pour  $i \neq i'$ . On a donc facilement la preuve que le problème d'interpolation polynomiale possède une solution unique dès que les points de mesures ont des abscisses distinctes deux à deux, puisque le problème se réduit à résoudre

$$QA = Y, (11)$$

qui a effectivement une solution et une seule si le déterminant de Q est non nul.

En fait, on peut dans ce cas écrire directement que la relation  $y_i = P(x_i)$  doit être vérifiée pour tous les i, ce qui revient à écrire QA = Y.

#### 3.1 Méthode de Lagrange.

On peut trouver une solution au problème d'interpolation polynomiale en utilisant une astuce que l'on appelle généralement interpolation par la méthode de Lagrange. Cette méthode repose sur la remarque suivante : si un polynome s'annule en  $x_i$ , alors il est divisible par  $X - x_i$ . On peut donc construire facilement une solution a tous les problèmes d'interpolation polynomiale dont le second membre Y dans l'équation (11) est un vecteur qui a toutes ses composantes nulles sauf une. En effet, si on considère un ensemble de mesures représenté par les couples  $(x_i, y_i)$ , avec les  $x_i$  distincts 2 à 2, on peut trouver un polynôme  $\mathcal{L}_i(X)$  tel que :

$$\mathcal{L}_i(x_i) = \delta_i^j$$
.

Autrement dit, on peut résoudre l'équation

$$QA = \begin{pmatrix} 0 \\ \vdots \\ y_j = 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Pour celà, il suffit d'utiliser le polynome d'interpolation

$$\mathcal{L}_{j}(X) = \frac{(X - x_{0}) \cdots (X - x_{j} - 1)(X - x_{j+1}) \cdots (X - x_{N-1})}{(x_{j} - x_{0}) \cdots (x_{j} - x_{j} - 1)(x_{j} - x_{j+1}) \cdots (x_{j} - x_{N-1})} = \frac{\prod_{i \neq j} (X - x_{i})}{\prod_{i \neq j} (x_{j} - x_{i})}$$

On vérifie facilement que

$$\mathcal{L}_j(x_j) = \frac{\prod_{i \neq j} (x_j - x_i)}{\prod_{i \neq j} (x_j - x_i)} = 1$$

et que  $\mathcal{L}_j(x_k) = 0$  si  $k \neq j$ .

Le polynôme d'interpolation que l'on cherchait au départ est alors obtenu en superposant les polynomes  $y_j$ . $\mathcal{L}_j$ . En effet en posant

$$P(X) = \sum_{j=0}^{N-1} y_j \mathcal{L}_j(X) ,$$

on obtient le polynôme d'interpolation P, puisque

$$P(x_i) = \sum_{j=0}^{N-1} y_j \mathcal{L}_j(x_i) = \sum_{j=0}^{N-1} y_j \delta_i^j = y_i.$$

Ceci permet également de trouver la matrice inverse de la matrice de Van Der Monde Q associée aux points  $(x_i)_{i=0..(N-1)}$ . En effet, le vecteur  $A_j$  formé par les coefficients de  $\mathcal{L}_j$  est la solution de  $(QA_j)_i = \delta_i^j$ , il représente donc le vecteur de la colonne j de la matrice inverse de Q.

#### 3.2 Exemples.

#### 3.2.1 Echantillonage régulier.

On va illustrer la méthode d'interpolation polynomiale en essayant de construire une approximation de la fonction  $f(x) = 1/(1+x^2)$  sur l'intervalle [-3,3]. Pour celà, on choisit N points au hasard dans l'intervalle, et on construit des "mesures" sous la forme de couples  $(x_i, y_i = 1/(1+x_i^2))$ . On utilise ensuite la méthode de Lagrange pour construire le polynôme d'interpolation que l'on va utiliser comme approximation de la fonction f.

Dans un premier temps, on va choisir des abscisses  $(x_i)$  réparties régulièrement dans l'intervalle. Un petit calcul avec octave permet de réaliser ces opérations, on a programmé la procédure d'interpolation et on obtient les résultats suivants pour 10 et 30 points :

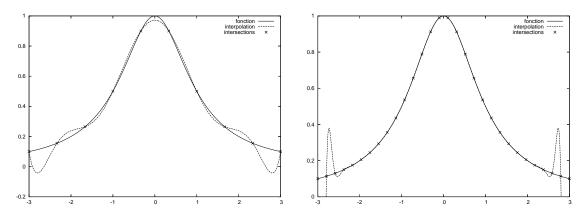


Fig. 3 – Interpolation polynomiale d'une fonction avec un échantillonnage uniforme sur l'intervalle pour 10 et 30 points.

On constate sur la figure que la fonction est plutôt bien approchée sur la majeure partie de l'intervalle, mais que des problèmes se présentent au bord. Par ailleurs, contrairement à ce qu'on pourrait supposer a priori, l'augmentation du nombre de points d'interpolation n'améliore pas le résultat et c'est même le contraire qui se produit, puisque les oscillations en bord d'intervalle augmentent fortement en passant de 10 à 30 points (au point que l'on a été obligé de restreindre l'étalement en y de la représentation pour continuer à voir la forme de la courbe de départ).

#### 3.2.2 Points de Gauss.

On peut recommencer l'opération précédente en utilisant d'autres abscisses pour l'interpolation. En essayant de concentrer les points aux endroits où l'interpolation fonctionne moins bien, on devrait pouvoir réduire l'erreur commise. On va utiliser les points  $x_i = 3 * cos((2*i+1)\pi/(2N))$  qui sont beaucoup plus concentrés autour de -3 et 3. Le résultat obtenu est le suivant : Il est évident sur la figure que les résultats sont globalement

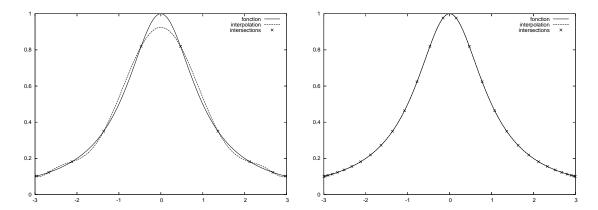


Fig. 4 – Interpolation polynomiale d'une fonction avec un échantillonnage concentré aux bords de l'intervalle pour 10 et 30 points.

meilleurs, en particulier parce qu'ils s'améliorent lorsque le nombre de points augmente. De plus il n'y a plus d'oscillations de grande amplitude ce qui implique une approximation plus uniforme de la fonction.