



Institut
Mines-Telecom

Quantification Scalaire et Prédictive

Marco Cagnazzo,
cagnazzo@telecom-paristech.fr

SIGMA 201



Plan

Introduction

Quantification uniforme

Quantification optimale

Quantification scalaire prédictive

Plan

Introduction

Quantification uniforme

Quantification optimale

Quantification scalaire prédictive

Définitions

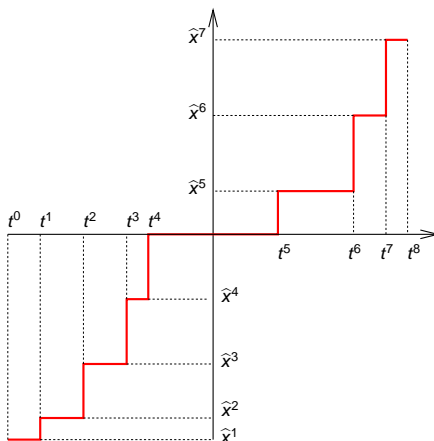
$$Q : x \in \mathbb{R} \rightarrow y \in \mathcal{C} = \{\hat{x}^1, \hat{x}^2, \dots, \hat{x}^L\} \subset \mathbb{R}$$

- ▶ \mathcal{C} : Dictionnaire, c'est un sous-ensemble discret de \mathbb{R}
- ▶ \hat{x}^i : niveau de quantification, niveau de restitution, codeword, mot de code
- ▶ $e = x - Q(x)$: Bruit de quantification
- ▶ $\Theta^i = \{x : Q(x) = \hat{x}^i\}$: Régions de décision

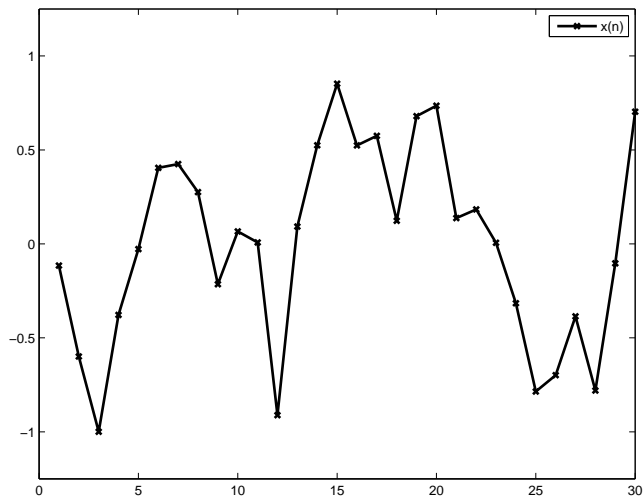
Un quantificateur scalaire (QS) est complètement défini par ses régions et ses niveaux

Définition : Quantification scalaire (QS)

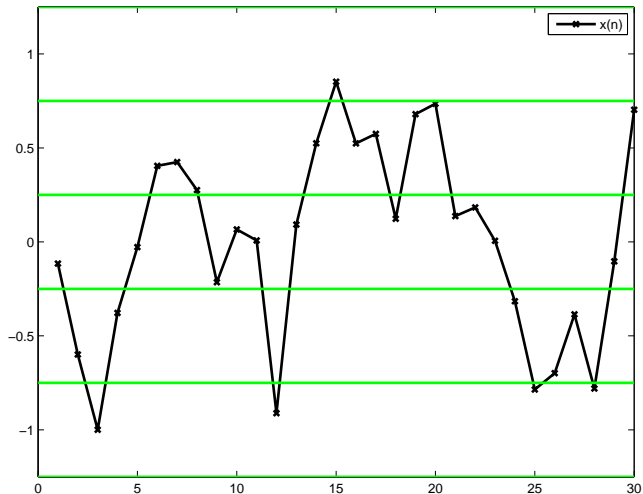
$$Q : x \in \mathbb{R} \rightarrow y \in \mathcal{C} = \{\hat{x}^1, \hat{x}^2, \dots, \hat{x}^L\} \subset \mathbb{R}$$



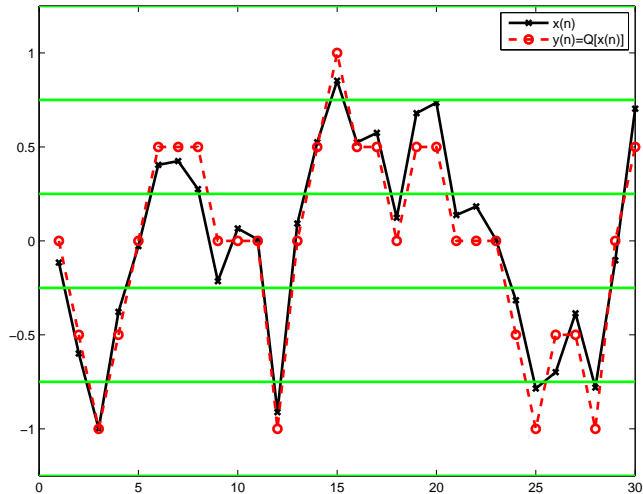
Exemple 1



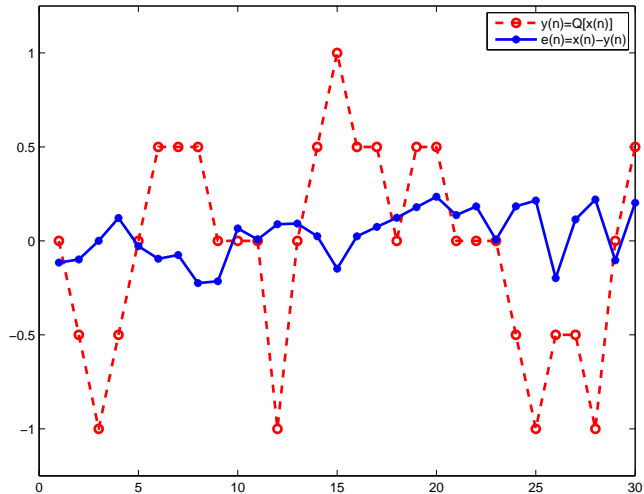
Exemple 1



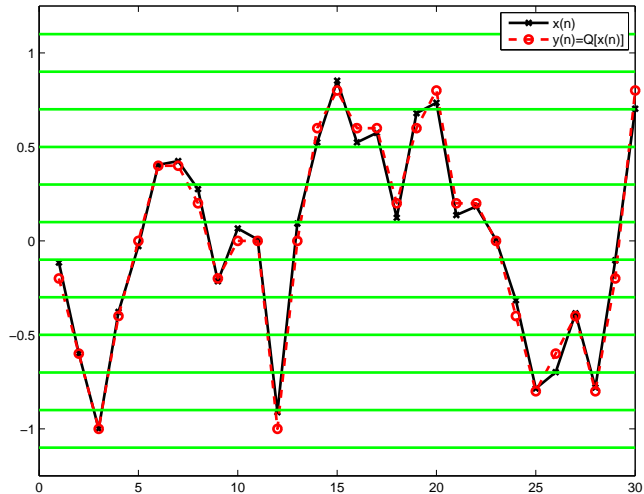
Exemple 1



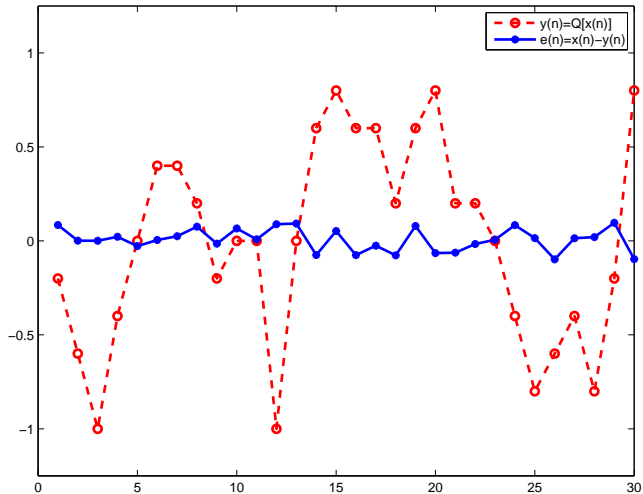
Exemple 1



Exemple 2



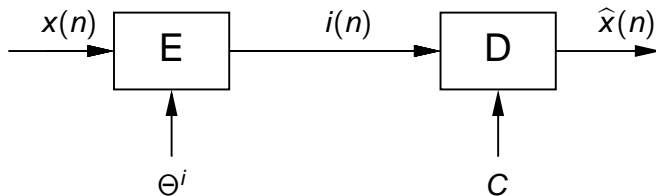
Exemple 2



Quantification d'un signal discret

- ▶ Typiquement on quantifie une suite (un signal discret) avec un QS régulier
- ▶ Les signaux d'intérêt ont une dynamique limitée $[-A, A]$
- ▶ Dans ce cas, $t^0 = -A$ et $t^L = +A$, toutes les régions sont finies, et l'erreur maximum est bornée
- ▶ La quantification consiste donc à associer à chaque échantillon d'entrée $x(n)$ un indice $i(n)$, qui détermine la région de décision, c.-à.-d. le mot de code $\hat{x}^{i(n)}$

Quantification comme codage/décodage



- ▶ Ce que le codeur envoie c'est les $i(n)$
- ▶ Le "décodage" associe à $i(n)$ le mot de code $\hat{x}^{i(n)}$
- ▶ Souvent (avec abus de langage) on appelle quantification l'application $x \rightarrow i$ et *quantification inverse* le décodage $i \rightarrow \hat{x}^i$

Débit d'un QS

- ▶ Le débit d'un QS est le nombre de bit nécessaire pour représenter les indices $i(n)$
- ▶ Par définition, $R = \log_2 L$
- ▶ Cela correspond à un codeur à longueur fixe et à un nombre de niveaux que soit une puissance entière de 2
- ▶ Il est possible de réduire ce débit avec le *codage sans pertes*

Distorsion

- ▶ On définit la distorsion ponctuelle comme l'erreur quadratique :

$$d[x(n), \hat{x}(n)] = |e(n)|^2 = |x(n) - \hat{x}(n)|^2$$

- ▶ Si on considère tout le signal $x(\cdot)$ de durée N , on utilise comme distorsion l'erreur quadratique moyenne :

$$D = \frac{1}{N} \sum_{n=0}^{N-1} d[x(n), \hat{x}(n)]$$

Distorsion : cas aléatoire

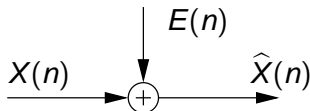
- Souvent on utilise des modèles aléatoires **centrés** pour les signaux. Dans ce cas, l'EQM associé à la QS du processus aléatoire X est :

$$D = E \left\{ |X(n) - Q(X(n))|^2 \right\} = E \left\{ |E(n)|^2 \right\}$$

- La distorsion est donc la puissance du processus aléatoire $E(n) = X(n) - Q(X(n))$
- On indiquera la distorsion d'un QS comme σ_Q^2

Distorsion : modèles pour le signal de bruit

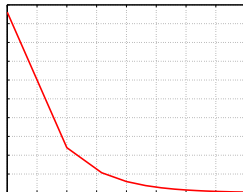
- ▶ Un modèle commun pour la quantification est le modèle additif



- ▶ Le bruit de quantification est donc vu comme un bruit additif
- ▶ $E(n)$ et $X(n)$ sont modélisés comme incorrélés
- ▶ $E(n)$ est modélisé comme uniforme
- ▶ E est un processus blanc (corrélation impulsionnelle)

Courbe débit/distorsion d'un QS

- ▶ Souvent on caractérise un QS par rapport au nombre de niveaux L
- ▶ Le débit croît avec L : $R = \log_2 L$
- ▶ Pour tout les cas d'intérêt, la distorsion décroît avec L , mais la relation explicite entre D et L est plus difficile à déterminer
- ▶ En général, un QS est donc caractérisé par une courbe paramétrique $R(L), D(L)$
- ▶ Il est intéressant de trouver la relation explicite entre D et R : c'est la courbe débit/distorsion $D = D(R)$





Plan

Introduction

Quantification uniforme

Quantification optimale

Quantification scalaire prédictive

Quantification uniforme

Un QS uniforme (QU) est caractérisé par :

- ▶ Régions de décision d'amplitude fixe : $\forall i, t^i = t^{i-1} + \Delta$
- ▶ Les niveaux de restitution sont les centres des régions :
$$\hat{x}^i = \frac{t^i + t^{i-1}}{2}$$

Le QU est simple, minimise l'erreur maximale et est optimale pour des v.a. uniformes.

On a encore :

- ▶ $\Delta^i = \Delta = 2A/L$

Quantification uniforme : calcul de la distorsion

Hypothèse : $X \sim \mathcal{U}(-A, A)$. Trouver $\sigma_Q^2 = \mathbb{E} \left[(X - \hat{X})^2 \right]$

Quantification uniforme : calcul de la distorsion

Hypothèse : $X \sim \mathcal{U}(-A, A)$. Trouver $\sigma_Q^2 = \mathbb{E}[(X - \hat{X})^2]$

$$\sigma_Q^2 = \mathbb{E}[(X - \hat{X})^2] = \int_{-A}^A p_X(u)[u - Q(u)]^2 du$$

Quantification uniforme : calcul de la distorsion

Hypothèse : $X \sim \mathcal{U}(-A, A)$. Trouver $\sigma_Q^2 = \mathbb{E}[(X - \hat{X})^2]$

$$\begin{aligned}\sigma_Q^2 &= \mathbb{E}[(X - \hat{X})^2] = \int_{-A}^A p_X(u)[u - Q(u)]^2 du \\ \dots &= \sum_{i=1}^L \int_{\Theta^i} \frac{1}{2A} [u - \hat{x}^i]^2 du = \frac{1}{2A} \sum_{i=1}^L \int_{\hat{x}^i - \Delta/2}^{\hat{x}^i + \Delta/2} [u - \hat{x}^i]^2 du \\ &= \frac{1}{2A} \sum_{i=1}^L \int_{-\Delta/2}^{\Delta/2} t^2 dt = \frac{1}{2A} L \frac{\Delta^3}{12} = \frac{\Delta^2}{12}\end{aligned}$$

En effet le bruit de quantification dans ce cas est une v.a. uniforme en $(-\Delta/2, \Delta/2)$

Quantification uniforme : courbe D(R)

$$D = \frac{\Delta^2}{12} = \frac{4A^2}{12L^2} = \frac{A^2}{3 \cdot 2^{2R}} = \sigma_X^2 2^{-2R}$$

On peut mesurer la qualité par le rapport signal sur bruit :

$$\begin{aligned} \text{SNR} &= 10 \log_{10} \frac{E \{X^2\}}{D} = 10 \log_{10} \frac{\sigma_X^2}{\sigma_X^2 2^{-2R}} \\ &= 10 \log_{10} 2^{2R} \approx 6R \end{aligned}$$

Quantification uniforme en haute résolution

- ▶ Hypothèse : $L \rightarrow +\infty$, X v.a. quelconque
- ▶ En HR, pour tout Θ^i on approxime p_X comme constante.
- ▶ Donc le bruit de quantification en Θ^i est $\mathcal{U}(-\frac{\Delta}{2}, \frac{\Delta}{2})$
- ▶ Pour la probabilité totale, $E \sim \mathcal{U}(-\frac{\Delta}{2}, \frac{\Delta}{2})$
- ▶ Donc :

$$D = \frac{\Delta^2}{12} = \frac{A^2}{3} 2^{-2R}$$

Quantification uniforme en haute résolution

On peut écrire :

$$\text{SNR} = 10 \log_{10} \frac{\mathbb{E}\{X^2\}}{D} = 10 \log_{10} \frac{\sigma_X^2}{A^2/3} 2^{2R} \approx 6R - 10 \log_{10} \frac{\gamma^2}{3}$$

où on a défini $\gamma^2 = \frac{X_{\max}^2}{\sigma_X^2} = \frac{A^2}{\sigma_X^2}$ rapport entre puissance de crête et puissance moyenne (facteur de charge)

On trouve encore :

$$D = \frac{A^2}{3} 2^{-2R} = \frac{\gamma^2}{3} \sigma_X^2 2^{-2R} = K_X \sigma_X^2 2^{-2R}$$

Le facteur de charge est donc un paramètre unique qui caractérise les performance de la QU en HR.

Quantification scalaire : exemple sur image couleurs

Image Originale, 24 bpp



Quantification scalaire : exemple sur image couleurs

Débit 21 bpp PSNR 47.19 dB TC 1.143



Quantification scalaire : exemple sur image couleurs

Débit 18 bpp PSNR 42.38 dB TC 1.333



Quantification scalaire : exemple sur image couleurs

Débit 15 bpp PSNR 36.97 dB TC 1.600



Quantification scalaire : exemple sur image couleurs

Débit 12 bpp PSNR 31.40 dB TC 2.000



Quantification scalaire : exemple sur image couleurs

Débit 9 bpp PSNR 29.26 dB TC 2.667



Quantification scalaire : exemple sur image couleurs

Débit 6 bpp PSNR 27.83 dB TC 4.000



Quantification scalaire : exemple sur image couleurs

Débit 3 bpp PSNR 25.75 dB TC 8.000





Plan

Introduction

Quantification uniforme

Quantification optimale

Quantification scalaire prédictive

Quantification optimale

Pour une densité de probabilité $p_X(x)$ donnée, déterminer le quantificateur qui minimise la distorsion pour un débit donné. Problème équivalent à déterminer les seuils t^i et les niveaux \hat{x}^i .
Solutions :

- ▶ Solution analytique en *haute résolution*: $D = c_X \sigma_X^2 2^{-2R}$
- ▶ Si l'hypothèse de haute résolution n'est pas satisfaite, on peut atteindre un minimum local de la distorsion avec l'algorithme de Max-Lloyd

Quantification optimale en haute résolution

Si on définit $U = X/\sigma_X$, on peut écrire :

$$\sigma_Q^2 = c_X \sigma_X^2 2^{-2R} \quad \text{avec}$$

$$c_X = \frac{1}{12} \left[\int_{\mathbb{R}} p_U^{1/3}(t) dt \right]^3$$

c_X est appelé facteur de forme, car il dépend seulement de la forme de la PDF de X (c.-à.-d. de p_U) et pas de sa variance

Facteurs de forme pour PDF communes :

	Uniforme	Gaussienne
c_X	1	$\frac{\sqrt{3}}{2}\pi \approx 2.72$

Quantification optimale non-HR

- ▶ Les hypothèses d'haute résolution ne sont pas toujours respectées
- ▶ Les cas plus intéressants sont au contraire ceux où le débit disponible est faible
- ▶ **On dispose pas de formules pour la quantification optimale à faible débit**
- ▶ En revanche, on peut trouver des conditions nécessaires et pour avoir un QO
- ▶ Ces conditions permettent de déterminer un algorithme de quantification (algorithme de Lloyd-Max)

Quantification optimale

Algorithme de Lloyd-Max

1. initialiser les régions (p.e. uniforme)
2. trouver les *meilleures* régions pour le dictionnaire donné

$$t^i = \frac{\hat{x}^i + \hat{x}^{i+1}}{2}, \quad i \in \{1, \dots, L-1\}$$

3. trouver le *meilleur* dictionnaire pour les régions données

$$\hat{x}^i = E[X|X \in \Theta_i] = \frac{\int_{\Theta_i} x p_X(x) dx}{\int_{\Theta_i} p_X(x) dx}$$

4. boucler en 2 jusqu'à la convergence

Algorithme de Lloyd-Max pour les données

- ▶ On a pas toujours les distributions de probabilités des signaux, souvent on a seulement un ensemble de M observations
- ▶ L'algorithme de Lloyd-Max se modifie comme il suit :
 1. Soit $\mathcal{X} = \{u_1, u_2, \dots, u_M\}$ l'ensemble des données à quantifier
 2. Initialisation ($k=0$) avec un dictionnaire quelconque (p.e. uniforme) : $\mathcal{C}^{(k)} = \{\hat{x}_0^i\}_{i=1, \dots, L}$
 3. Règle du plus proche voisin :

$$W_k^i = \{u_m \in \mathcal{X} : \forall j \neq i \|u_m - \hat{x}_k^i\| \leq \|u_m - \hat{x}_k^j\|\}$$

4. Règle du centroïde : $\hat{x}_{k+1}^i = \frac{1}{|W_k^i|} \sum_{u_m \in W_k^i} u_m$
5. Boucler en 3 jusqu'à convergence

Convergence de l'algorithme de Lloyd-Max

- ▶ Rien n'assure la convergence à l'optimum global
- ▶ Le résultat dépend de l'**initialisation**
- ▶ En tout cas, à chaque itération la distorsion n'augmente pas
- ▶ Critère d'arrêt :
 - ▶ Soit $D^{(k)}$ la distorsion du k -ème quantificateur
 - ▶ Une condition d'arrêt typique est :

$$\frac{D^{(k)} - D^{(k+1)}}{D^{(k)}} \leq \epsilon$$



Algorithme de Lloyd-Max

Bilan de la quantification scalaire

- ▶ D(R), QU et v.a. uniforme : $D = \sigma_X^2 2^{-2R}$
- ▶ D(R), QU et v.a. non uniforme en HR : $D = K_X \sigma_X^2 2^{-2R}$
- ▶ D(R), QO et v.a. non uniforme en HR : $\sigma_Q^2 = c_X \sigma_X^2 2^{-2R}$
- ▶ c_X est une constante qui dépend de la distribution de X (facteur de forme)
 - ▶ $c_X = 1$ dans le cas uniforme
 - ▶ $c_X = \frac{\sqrt{3}}{2} \pi$ dans le cas gaussien
- ▶ À faible resolution, LM produit un QNU localement optimale



Plan

Introduction

Quantification uniforme

Quantification optimale

Quantification scalaire prédictive

Codage prédictive

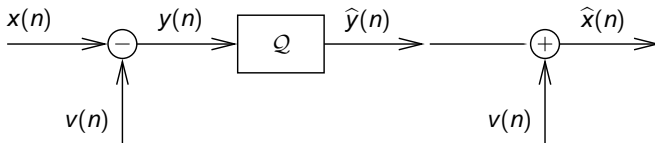
Principes

- ▶ La seule quantification est peu efficace pour la compression
- ▶ Modèle sous-jacent trop simple : échantillons indépendants et tous également importants
- ▶ Idée : exploiter la corrélation entre échantillon par une *prédiction*
- ▶ Réduction de la variance

Schéma de codage

Schéma en boucle ouverte

- ▶ L'échantillon $x(n)$ est lié aux échantillons passés (et futurs)
- ▶ On utilise les voisins de $x(n)$ pour le prédire
- ▶ Si on fait une bonne prédiction, $v(n) \approx x(n)$



- ▶ Comment on fait la prédiction ?
- ▶ Qu'est-ce qu'on gagne ?

Gain de prédiction

Erreur sur la prédiction = erreur sur le signal :

$$q(n) = y(n) - \hat{y}(n) = x(n) - v(n) - \hat{x}(n) + v(n) = \bar{q}(n)$$

Donc l'objectif de la QS prédictive devient celui de minimiser la distorsion de y

Gain de codage :

$$\text{SNR}_p = 10 \log_{10} \frac{\sigma_X^2}{D} = 10 \log_{10} \frac{\sigma_X^2}{\sigma_Y^2} + 10 \log_{10} \frac{\sigma_Y^2}{D} = G_P + G_Q$$

La prédiction doit produire un signal d'erreur dont la variance est inférieure à la variance du signal d'origine

Exemple

$$X(n) \sim \mathcal{N}(0, \sigma^2)$$

$$V(n) = X(n-1)$$

$$\mathbb{E}[X(n)X(m)] = \sigma^2 \rho^{|n-m|}$$

$$\rho : G_P > 0 ?$$

Exemple

$$\begin{aligned} X(n) &\sim \mathcal{N}(0, \sigma^2) & \mathbb{E}[X(n)X(m)] &= \sigma^2 \rho^{|n-m|} \\ V(n) &= X(n-1) & \rho : G_P &> 0 ? \end{aligned}$$

$Y(n) = X(n) - X(n-1)$ v.a. Gaussienne centrée

$$\sigma_Y^2 = \mathbb{E}[(X(n) - X(n-1))^2] = 2\sigma^2 - 2\sigma^2\rho$$

$$G_P = 10 \log_{10} \frac{\sigma_X^2}{\sigma_Y^2} = 10 \log_{10} \frac{\sigma^2}{2(1-\rho)\sigma^2}$$

$$G_P > 0 \Leftrightarrow \rho > \frac{1}{2}$$

Prédicteurs

- ▶ On s'intéressera aux prédicteurs linéaires : simples et optimaux dans le cas Gaussien

$$v(n) = - \sum_{i=1}^P a_i x_{n-i} \quad \text{Filtre de prédiction à } P \text{ paramètres}$$

$$y(n) = x(n) - v(n) = \sum_{i=0}^P a_i x_{n-i} \quad \text{Erreur de prédiction}$$

- ▶ avec $a_0 = 1$.
- ▶ On peut donc voir y comme filtrage de x avec le filtre de fonction de transfert :

$$A(z) = 1 + a_1 z^{-1} + \dots + a_P z^{-P}$$

- ▶ Filtre optimale : minimisation de σ_Y^2

Modèle AR du signal

- ▶ Si $Y(z) = A(z)X(z)$, $X(z) = \frac{Y(z)}{A(z)}$
- ▶ On montre que, si la prédiction est optimale, $Y(z)$ est du bruit blanc de puissance σ_Y^2
- ▶ Dans ce cas, la DSP de X est $S_X(f) = \frac{\sigma_Y^2}{|A(f)|^2}$
- ▶ Le modèle sous-jacente pour X est celui d'un signal auto-regressif :

$$X(z) = \frac{Y(z)}{1 + a_1 z^{-1} + \dots + a_P z^{-P}} \quad \Leftrightarrow$$
$$x(n) + a_1 x(n-1) + \dots + a_P x(n-P) = y(n)$$

- ▶ $x(n)$ est donc le filtrage AR de $y(n)$

Choix du prédicteur

Problème :

Trouver le vecteur (filtre linéaire) \underline{a} qui minimise :

$$\sigma_Y^2 = E \left\{ Y^2(n) \right\} = E \left\{ \left[X(n) + \sum_{i=1}^P a_i X(n-i) \right]^2 \right\}$$

Choix du prédicteur

$$\begin{aligned}\sigma_Y^2 &= \mathbb{E}\{X^2(n)\} + 2 \sum_{i=1}^P a_i \mathbb{E}\{X(n)X(n-i)\} + \sum_{i=1}^P \sum_{j=1}^P a_i a_j \mathbb{E}\{X(n-i)X(n-j)\} \\ &= \sigma_X^2 + 2 \underline{r}_X^t \underline{a} + \underline{a}^t \mathbf{R}_X \underline{a}\end{aligned}$$

avec :

$$\underline{r}_X = [r_X(1) \dots r_X(P)] \quad \mathbf{R}_X = \begin{bmatrix} r_X(0) & r_X(1) & r_X(2) & \dots & r_X(P-1) \\ r_X(1) & r_X(0) & r_X(1) & \dots & r_X(P-2) \\ r_X(2) & r_X(1) & r_X(0) & \dots & r_X(P-3) \\ \dots & \dots & \dots & \dots & \dots \\ r_X(P-2) & r_X(P-3) & r_X(P-4) & \dots & r_X(1) \\ r_X(P-1) & r_X(P-2) & r_X(P-3) & \dots & r_X(0) \end{bmatrix}$$

$$r_X(k) = \mathbb{E}\{X(n)X(n-k)\}$$

Choix du prédicteur

Minimisation de la variance de Y :

$$\frac{\partial \sigma_Y^2}{\partial \underline{a}} = 2\underline{r} + 2\mathbf{R}_X \underline{a} = 0$$

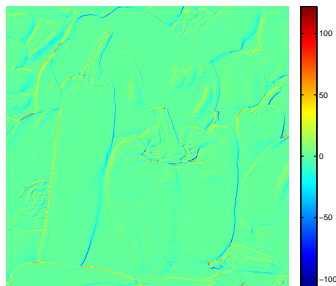
On a donc :

$$\underline{a}^{\text{opt}} = -\mathbf{R}_X^{-1} \underline{r} \qquad \sigma_Y^2 = \sigma_X^2 + \underline{r}^t \underline{a}^{\text{opt}}$$

L'autocorrélation r_X peut généralement être estimé avec :

$$\hat{r}_X(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} X(n)(X(n+k))$$

Méthodes prédictives : exemple



$$\sigma_X^2 = 2903$$

Prédicteur :

	b	
a	x	

$$v_{n,m} = ax_{n-1,m} + bx_{n,m-1}$$

a	b	σ_Y^2	G_P
1/2	1/2	78.7	15.67 dB
0.45	0.55	78.4	15.68 dB

Prédicteur :

c	b	d
a	x	

$$v_{n,m} = ax_{n-1,m} + bx_{n,m-1} + cx_{n-1,m-1} + dx_{n-1,m+1}$$

a	b	c	d	σ_Y^2	G_P
0.66	0.72	-0.39	0	70.56	16.14 dB
0.66	0.54	-0.38	0.18	62.12	16.70 dB