Méthodes d'optimisations

Loic Huguel

7 février 2017

Table des matières

1 Formulation

1.1 Cas général :

Soit le système $Y = f(U, X)$ avec f une fonction de $\Re^m \rightarrow \Re^n$, X vecteur d'état et (Ui, Yi) k couples d'entrées/sorties obtenus par l'expérience. On cherche alors quel est le vecteur X qui permet de décrire au mieux les couples de données (Ui, Yi), au sens d'un certain critère S, que on cherchera à minimiser.

Typiquement on choisi le critère des moindres carrés :

$$S(X) = \sum_{i=0}^{k}(Yi - f(Ui, X))^2$$

Sous forme vectorielle on notera :

$$S(X) = |Y - F(U, X)|$$

Avec

$$U = (U_0, .., U_k) \quad et \quad Y = (Y_0, .., Y_k)^T$$

La solution est donc :

$$X = argmin(S(X))$$

1.2 Cas linéaire :

Dans le cas ou la fonction f est linéaire par rapport au vecteur d'état X. On peut alors écrire le système d'equations suivant :

$$UX = Y$$

Le problème revient alors à calculer la pseudo-inverse de la matrice U. La solution est donc :

$$X = U^{-1}Y$$

Ou U^{-1} est la pseudo inverse de U. Généralement on calcule U^{-1} au sens des moindres carrés, c'est à dire tel que $|Y - UX|$ soit minimale.

2 Méthodes Générales

2.1 Descente de gradient

On choisit un point de départ X_0 puis on cherche un état suivant X_{k+1} tel que $S(X_{k+1}) < S(X_k)$ Pour ce faire, on cherche dans quelle direction δ la pente est minimum lorsque on est en X_k.

$$f(Ui, X_k + \delta) = f(Ui, X_k) + \delta J(Ui, X_k)$$

Ou J est le Jacobien de f en X_k.

$$S(X_k + \delta) = \sum_{i=0}^{k}(Yi - f(Ui, X_k) - \delta J(Ui, X_k))^2$$

$$\frac{dS(X_k + \delta)}{d\delta} = \sum_{i=0}^{k} 2J(Ui, X_k)(Yi - f(Ui, X_k) - \delta J(Ui, X_k)$$

Puis si on pose $\frac{dS(X_k+\delta)}{d\delta} = 0$ en vectoriel cela donne :

$$J^T(Y - F(U, X_k) - \delta J) = 0$$

$$\delta = -(J^T J)^{-1} J^T (Y - F(U, X_k))$$

$$X_{k+1} = X_k + \delta$$

2.2 Méthode de Newton

2.3 Levenberg-Maquardt

La technique de Levenberg-Maquardt reprend l'équation de la technique du gradient en amortissant par un paramètre λ.

$$\delta = -(J^T J - \lambda I)^{-1} J^T (Y - F(U, X_k))$$

Maquardt suggère alors de pondérer la matrice identité par la la diagonale de $J^T J$:

$$\delta = -(J^T J - \lambda diag(J^T J))^{-1} J^T (Y - F(U, X_k))$$

Le paramètre λ est alors ajusté en fonction de la rapidité de la convergence. Si on converge rapidement alors on diminue λ et réciproquement.
Cependant la notion de rapidité de convergence est très subjective, car elle dépend de l'échelle à laquelle on regarde le système.

3 Méthodes Linéaires

3.1 Décomposition en valeurs singulières SVD

La décomposition en valeur singulière consiste à décomposer une matrice de la manière suivante :

$$A = UWV^*$$

U et V des matrices unitaires c'est à dire que $UU^* = U^*U = I$
L'opération "*" est le conjugué Hermitien. Si U est une matrice à coefficients réels $U^* = U^T$
Cette décomposition permet de calculer une pseudo inverse au sens des moindre carrés, de la matrice M.

$$A^{-1} = VW^{-1}U^*$$

en effet on a bien avec les propriétés des matrices unitaires :

$$A^{-1}A = (VW^{-1}U^*)(UWV^*) = I$$

$$AA^{-1} = (UWV^*)(VW^{-1}U^*) = I$$