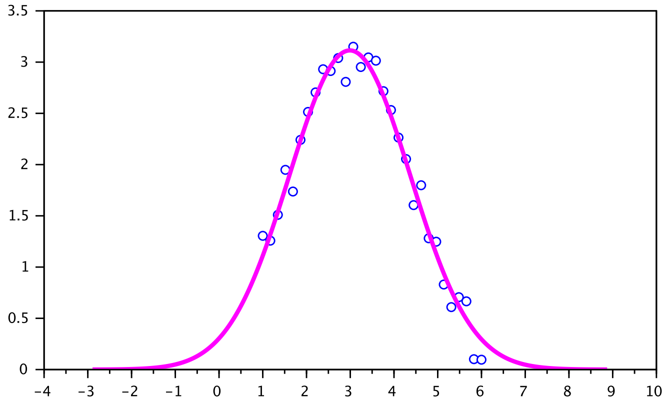
MT94/P24/TD - Problèmes de moindres carrés non-linéaires



On considère m couples $(t_i, y_i)_{i=1\ldots m}$. Les vecteurs `t` et `y` sont dans le fichier `dataGaussian.sod` disponible sur Moodle. On cherche à approcher au mieux ces données à l'aide de la fonction

$$f(t) = a \exp\left(-(t-\mu)^2/\sigma^2\right),$$

au sens des moindres carrés, c'est à dire de manière à ce que la quantité

$$E(a,\mu,\sigma) = \sum_{i=1}^{m} (f(t_i) - y_i)^2,$$

soit minimale.

1 Résolution approchée avec le « log trick »

Le « log trick » consiste à chercher le minimum de la fonction

$$E_{\log}(a,\mu,\sigma) = \sum_{i=1}^{m} (\log f(t_i) - \log y_i)^2.$$

C'est un problème différent du problème initial, mais qui est plus facile à résoudre (et qui peut donner une solution approchée).

1. Montrez que
$$\log f(t) = x_1 + x_2 t + x_3 t^2,$$
où les composantes du vecteur $\mathbf{x}$ s'écrivent en fonction de a, μ, σ.
2. Montrez qu'il existe une matrice $\mathbf{A}$ et un vecteur $\mathbf{b}$ tels que
$$E_{\log}(a,\mu,\sigma) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2.$$
En déduire que les valeurs de a, μ, σ minimisant $E_{\log}(a,\mu,\sigma)$ peuvent être calculées à partir de la solution d'un problème de moindres carrés linéaire.
3. Résolvez ce problème avec Scilab et évaluez la qualité de l'approximation en représentant à l'écran les points $(t_i, y_i)_{i=1\ldots m}$ et la Gaussienne obtenue.

2 Résolution exacte

On s'intéresse maintenant à la résolution du problème initial, c'est à dire en cherchant le minimum de

$$E(a,\mu,\sigma) = \sum_{i=1}^{m} (f(t_i) - y_i)^2.$$

Pour utiliser la méthode de Levenberg-Marquard, on a vu en cours qu'il fallait au préalable calculer l'expression de la Jacobienne du vecteur résidu $\mathbf{r} \in \mathbb{R}^n$ défini par

$$r_i = f(t_i) - y_i,\ i = 1 \dots n,$$

par rapport aux paramètres inconnus.

1. Calculer la Jacobienne du vecteur résidu $\mathbf{r}$ par rapport au vecteur d'inconnues.
2. Résolvez le problème de moindres carrés avec la méthode de Levenberg-Marquardt. Utilisez dans un premier temps des valeurs initiales arbitraires $(1,1,1)$ et une valeur du paramètre de la méthode $\lambda = 0$. Ajustez λ pour faire converger la méthode (ou la faire converger plus rapidement).
3. Evaluez la qualité de l'approximation en représentant à l'écran les points $(t_i, y_i)_{i=1\dots m}$ et la Gaussienne obtenue. Que constatez-vous par rapport à l'approximation faite avec la méthode précédente ?
4. Que remarquez-vous sur le nombre d'itérations si les valeurs initiales des paramètres sont celles obtenues avec le « log trick » ? Déterminer dans ce cas la valeur minimale de λ permettant de garantir la convergence. Conclure.
5. Quel avantage éventuel voyez-vous à considérer (x_1, x_2, x_3) comme vecteur d'inconnues au lieu de (a,μ,σ) ? Comparer les résultats obtenus pour déterminer quel est le meilleur choix.