

TP : recalage d'un réseau routier

Yann Méneroux

13 Novembre 2020

L'objectif de ces travaux pratiques consiste à prétraiter des données géographiques en vue de l'analyse d'une trajectoire GPS issue d'un véhicule. On dispose des données suivantes :

- 1 fichier texte `gps.dat` contenant la séquence des positions (ordonnées chronologiquement) observées par un récepteur GPS embarqué dans un véhicule. Chaque point est exprimé par ses coordonnées planes $(X, Y) \in \mathbb{R}^2$ dans la projection cartographique Lambert 93 (système géodésique RGF93).
- 2 fichiers texte contenant le réseau routier aux alentours de l'Ile de la Cité (Paris), référencé dans une projection cartographique locale inconnue :
 - `vertex.dat`, repertoriant pour chaque sommet v_i (indice i du point en première colonne), ses coordonnées planes $(X_i, Y_i) \in \mathbb{R}^2$
 - `edge.dat`, la table topologique des arcs, repertoriant pour chaque arc e_{ij} du réseau (indice de l'arc en première colonne), les indices i et j (dans la table des sommets) des extrémités de l'arc et sa longueur métrique $\|e_{ij}\|$. On précise que les arcs sont non-orientés, c'est-à-dire que $e_{ij} = e_{ji}$.
- 1 fichier de script Matlab `convert.m` permettant de transformer les points saisis sur le Géoportail (cf question Q6).

Q1. Charger (sous forme matricielle à l'aide de la commande `load`) et représenter le réseau routier dans l'environnement Matlab. Les indices des sommets seront affichés à l'aide de l'instruction :

```
text(x, y, num2str(indice), 'FontSize', 4);
```

Q2. Charger et représenter en rouge la trace GPS. Expliquer le résultat graphique obtenu.

Q3. Dans ce TP, on va corriger les positions des sommets du réseau routier de sorte à les recalcr dans le système de coordonnées de la trace GPS. Justifier ce choix (*i.e.* pourquoi ne pas préférer à l'inverse adapter les coordonnées de la trace GPS dans le système du réseau routier?).

Q4. La transformation de coordonnées s'effectuant entre deux projections cartographiques, on utilisera la méthode des polynômes d'adaptation. Pour rappel, le passage des coordonnées d'un système $(X, Y) \in \mathbb{R}^2$ dans un second système $(X', Y') \in \mathbb{R}^2$ s'exprime par les relations :

$$X' = P_X(X) = \sum_{n=0}^p a_n X^n \quad Y' = P_Y(Y) = \sum_{n=0}^p b_n Y^n$$

avec P_X et $P_Y \in \mathbb{R}_p[X]$ deux polynômes de régression de degré $p \in \mathbb{N}$ et de coefficients (inconnus et à déterminer) : $\{a_n, b_n\}_{n=0..p}$. On rappelle qu'un point de liaison est un point dont on connaît les coordonnées (X, Y) et (X', Y') dans les deux systèmes. Déterminer (en fonction de p) le nombre minimum de points de liaison nécessaires à la résolution du système d'équations.

Dans la suite, on ne s'intéresse qu'à la classe des polynômes de degré $p = 3$. L'objectif des questions Q5 et Q6 va être de saisir ces points de liaison.

Q5. Pour relever un ensemble de points de liaison, on utilisera l'outil *Annoter la carte > Placer des points* du Géoportail (<https://www.geoportail.gouv.fr/>).

On rappelle que la qualité du recalage est croissante avec le nombre, l'homogénéité spatiale et la précision des points de liaison. Pour une bonne qualité de saisie, il est recommandé de sélectionner des points simples à apparier (sans ambiguïté, e.g. extrémités des ponts, intersections caractéristiques, décrochements de berge...) et bien résolus (éviter les grands carrefours, les routes à plusieurs voies...).

On commencera dans un premier temps par prendre une vingtaine de points (si les résultats obtenus à la question Q10 ne sont pas satisfaisants, revenir sur cette étape pour sélectionner plus de points). On pourra si besoin s'aider de l'annexe 1 pour établir plus facilement les correspondances entre les sommets du réseau et la photographie aérienne.

Exporter le fichier des points de liaison en format kml. Ouvrir le fichier avec un éditeur de texte et inspecter les coordonnées des points saisis.

De quel type de coordonnées s'agit-il (☐ cartésiennes ☐ géographiques ☐ planes)?

Q6. On souhaite exprimer ces coordonnées dans la même référence que le jeu de données GPS (RGF93 - Projection Lambert 93). Pour ce faire, on pourra utiliser au choix, un script Matlab ou bien un application en ligne.

a) Avec un script Matlab :

On pourra utiliser le script `convert_matlab.m` ou `convert_octave.m` (en fonction du logiciel utilisé) qui permet de lire le fichier kml de points saisis dans le Géoportail, et de le convertir en une matrice `links.dat` contenant les coordonnées (et indices) des points de liaison après projection en Lambert 93. Avant de lancer le script, on devra l'ouvrir avec un éditeur de texte pour modifier les adresse des fichier d'entrée/sortie. On pourra alors importer la matrice résultat en Matlab (sous le nom `L93`) avec la commande :

```
load links.dat.
```

b) Avec une application en ligne :

Avec cette option, il faut commencer par convertir le fichier kml généré à la question précédente en un fichier CSV comprenant 1 ligne par point de sélectionné et 3 colonnes séparées par une virgule :

```
id_du_point, longitude, latitude
```

Le fichier ne doit pas contenir d'entête.

Dans la barre du navigateur, entrer l'adresse suivante :

```
https://tool-online.com/conversion-coordonnees.php
```

Il faut impérativement effectuer les trois étapes ci-dessous dans l'ordre indiqué.

- Dans la fenêtre de droite, sélectionner la projection RGF93-Lambert-93 (pour avoir accès à cette projection, il faut auparavant remplacer le champ *WORLD* par *France*).
- Cliquer sur le bouton *Open text file* et sélectionner le fichier CSV (fichier des points de liaison en WGS84).
- Cliquer sur le lien *Download the results* et mettre le résultat sous un format compréhensible par Matlab.

Q7. Pour se prémunir du mauvais conditionnement (susceptible d'entraîner des problèmes de singularité des matrices à la précision machine), il est recommandé de normaliser l'ensemble des coordonnées des sommets du réseau routier (*i.e.* diviser les valeurs numériques X et Y des coordonnées de sommets issus du fichier `vertex.dat`, respectivement par les valeurs maximales X_{max} et Y_{max} de ces coordonnées).

Q8. On note $m \in \mathbb{N}$ le nombre de points de liaison saisis. Ecrire (sur papier) et implanter dans Matlab les matrices $A \in \mathbb{R}^{m \times 2(p+1)}$ et $B \in \mathbb{R}^m$ telles que le système de transformation présenté à la question Q4 puisse se mettre sous la forme linéaire :

$$\mathbf{A}\mathbf{X} = \mathbf{B}$$

avec $\mathbf{X} \in \mathbb{R}^{2(p+1)}$ le vecteur contenant les paramètres $\{a_n, b_n\}_{n=0..p}$ de la transformation. On pourra ordonner ces paramètres de sorte à rendre le plus simple possible le code Matlab de construction de la matrice \mathbf{A} .

Q9. On rappelle que sous l'hypothèse d'homogénéité de la précision de saisie des points de liaison, la solution du problème, obtenue par moindres carrés, s'exprime par :

$$\mathbf{X} = (\mathbf{A}^t \mathbf{A})^{-1} \mathbf{A}^t \mathbf{B}$$

Estimer les paramètres de transformation (ainsi que le RMSE associé¹) et représenter le nouveau réseau recalé par rapport à la trace GPS. L'erreur commise vous paraît-elle admissible ?

Note : si l'erreur RMSE est trop importante, on vérifiera que l'appariement réseau / orthophotographie des points de liaison a été correctement effectué (l'analyse des résidus doit généralement permettre de détecter la

1. La racine carrée de l'erreur moyenne quadratique.

présence de points de liaison aberrants).

Q10. Pour contrôler la qualité du recalage, on souhaite évaluer également la capacité de généralisation de la transformation, c'est-à-dire que l'on cherche à mesurer l'erreur commise en dehors du voisinage des points de liaison. La méthode traditionnellement employée (pour m points relevés) consiste à calculer les paramètres de transformation avec les $m - k$ premiers points de liaison, puis à évaluer l'erreur commise sur les k derniers points. On parle de validation croisée.

Implémenter une procédure de validation croisée pour mesurer la qualité du recalage (on prendra $k = \lceil \frac{m}{2} \rceil$).

