Matrix representation of the linear system encountered in the least-squares minimization

jerviedog

January 17, 2018

1 Introduction

This is a short summary of the derivation of the linear system encountered in the linear least-square minimization problem. Specifically, we want to adjust the curve $\boldsymbol{\mathcal{P}}_{+} = \sum_j \beta_j(x)(\boldsymbol{P}_j + \boldsymbol{D}_j)$ to minimize the distance from the point cloud to the curve. The displacements of the control points $\mathcal{D} = \{\boldsymbol{D}\}$ are the optimization variables. The result can also be used to the optimization problem of similar form.

2 TDM

The minimization problem is

$$\min_{\mathcal{D}} f(\mathcal{D}) = \min_{\mathcal{D}} \sum_k e_k(\mathcal{D}),$$

where

$$\begin{aligned} e_k(\mathcal{D}) &= e_{TD,k}(\mathcal{D}) \\ &= \frac{1}{2}\left[(\boldsymbol{\mathcal{P}}_{+}(t_k) - \boldsymbol{X}_k)^T \boldsymbol{N}_k\right]^2 \\ &= \frac{1}{2}\Big[(\sum_j \beta_{jk}(\boldsymbol{P}_j + \boldsymbol{D}_j) - \boldsymbol{X}_k)^T \boldsymbol{N}_k\Big]^2, \end{aligned}$$

and

$$\beta_{jk} = \beta_j(t_k).$$

In two-dimensional cases, the control point $\boldsymbol{P}_j = (P_{jx}, P_{jy})^T$, and the displacement of the control point $\boldsymbol{D}_j = (D_{jx}, D_{jy})^T$.

Suppose that there are n sample points in the point cloud, and m control points. Minimizing $f(\mathcal{D})$ leads to

$$\frac{\partial f}{\partial D_{ix}} = 0$$

and

$$\frac{\partial f}{\partial D_{iy}} = 0.$$

Because $f(\mathcal{D})$ is linear sum of squares, the minimization of $f(\mathcal{D})$ would give us a linear system $\mathbf{Ax} = \mathbf{b}$, the solution of which is the solution of the original optimization problem.

Below is the derivation of such linear system.

First, $f(\mathcal{D})$ can be rewritten as:

$$\begin{aligned}
f(\mathcal{D}) &= \sum_{k=1}^{n} e_k(\mathcal{D}) \\
&= \frac{1}{2}\sum_{k=1}^{n}\left[(\boldsymbol{\mathcal{P}}_{+}(t_k) - \boldsymbol{X}_k)^T \boldsymbol{N}_k\right]^2 \\
&= \frac{1}{2}\sum_{k=1}^{n}\left[(\boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k)^2 - 2(\boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k)(\boldsymbol{X}_k^T \boldsymbol{N}_k) + (\boldsymbol{X}_k^T \boldsymbol{N}_k)^2\right].
\end{aligned}$$

The derivative of $\boldsymbol{\mathcal{P}}_{+}(t_k)$ with respect to D_{ix} and D_{iy} are

$$\frac{\partial \boldsymbol{\mathcal{P}}_{+}(t_k)}{\partial D_{ix}} = \beta_{ik}\begin{pmatrix}1\\0\end{pmatrix},$$

and

$$\frac{\partial \boldsymbol{\mathcal{P}}_{+}(t_k)}{\partial D_{ix}} = \beta_{ik}\begin{pmatrix}0\\1\end{pmatrix}.$$

The derivative of $\boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k$ with respect to D_{ix} and D_{iy} are

$$\frac{\partial \boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k}{\partial D_{ix}} = \beta_{ik} N_{kx},$$

and

$$\frac{\partial \boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k}{\partial D_{iy}} = \beta_{ik} N_{ky}.$$

Then the derivative of f with respect to D_{ix} and D_{iy} are

$$\begin{aligned}
\frac{\partial f}{\partial D_{ix}} &= \sum_k \left[\beta_{ik} N_{kx} \boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k - \beta_{ik} N_{kx} (\boldsymbol{X}_k^T \boldsymbol{N}_k)\right] \\
&= \sum_k \left\{\beta_{ik} N_{kx}\Big[\sum_j \beta_{jk}(\boldsymbol{P}_j + \boldsymbol{D}_j)\Big]^T \boldsymbol{N}_k - \beta_{ik} N_{kx}(\boldsymbol{X}_k^T \boldsymbol{N}_k)\right\},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial f}{\partial D_{iy}} &= \sum_k \left[\beta_{ik} N_{ky} \boldsymbol{\mathcal{P}}_{+}(t_k)^T \boldsymbol{N}_k - \beta_{ik} N_{ky} (\boldsymbol{X}_k^T \boldsymbol{N}_k)\right] \\
&= \sum_k \left\{\beta_{ik} N_{ky}\Big[\sum_j \beta_{jk}(\boldsymbol{P}_j + \boldsymbol{D}_j)\Big]^T \boldsymbol{N}_k - \beta_{ik} N_{ky}(\boldsymbol{X}_k^T \boldsymbol{N}_k)\right\}.
\end{aligned}$$

$\partial f/\partial D_{ix} = 0$ result in the first m equations, while $\partial f/\partial D_{iy} = 0$ lead to the other m equations. We next would like to rewrite these equations in the form of $\mathbf{Ax} = \mathbf{b}$.

For the ith equation, that is, $\partial f/\partial D_{ix} = 0$,

$$\begin{aligned}
\text{LHS} &= \sum_k \left\{\beta_{ik} N_{kx}\Big[\sum_j (\beta_{jk}(D_{jx} N_{kx} + D_{jy} N_{ky}))\Big]\right\} \\
&= \sum_j \sum_k \beta_{ik}\beta_{jk} N_{kx} N_{kx} D_{jx} + \sum_j \sum_k \beta_{ik}\beta_{jk} N_{kx} N_{ky} D_{jy} \\
\text{RHS} &= \sum_k \left\{\beta_{ik} N_{kx}\big[(\boldsymbol{X}_k - \boldsymbol{\mathcal{P}}_k)^T \boldsymbol{N}_k\big]\right\}.
\end{aligned}$$

Note that here $\boldsymbol{\mathcal{P}}_k = \sum_j \beta_{jk} \boldsymbol{P}_j$.

For the $i+m$th equation, that is, $\partial f / \partial D_{iy} = 0$,

$$\begin{aligned}
\text{LHS} &= \sum_k \Big\{ \beta_{ik} N_{ky} \Big[\sum_j (\beta_{jk}(D_{jx} N_{kx} + D_{jy} N_{ky})) \Big] \Big\} \\
&= \sum_j \sum_k \beta_{ik} \beta_{jk} N_{kx} N_{ky} D_{jx} + \sum_j \sum_k \beta_{ik} \beta_{jk} N_{ky} N_{ky} D_{jy} \\
\text{RHS} &= \sum_k \Big\{ \beta_{ik} N_{ky} \big[(\boldsymbol{X}_k - \boldsymbol{\mathcal{P}}_k)^T \boldsymbol{N}_k \big] \Big\}.
\end{aligned}$$

Blah, blah, blah, at last, a matrix form $\mathbf{A}\mathbf{x} = \mathbf{b}$ is obtained:

$$\begin{aligned}
\text{LHS} &= \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^T \begin{pmatrix} \mathbf{N_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{N_y} \end{pmatrix} \begin{pmatrix} \mathbf{N_x} & \mathbf{N_y} \\ \mathbf{N_x} & \mathbf{N_y} \end{pmatrix} \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{D_x} \\ \mathbf{D_y} \end{pmatrix}, \\
\text{RHS} &= \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^T \begin{pmatrix} \mathbf{N_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{N_y} \end{pmatrix} \begin{pmatrix} \mathbf{F} \\ \mathbf{F} \end{pmatrix},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{B} &= \begin{pmatrix} \beta_{11} & \cdots & \beta_{m1} \\ \vdots & \ddots & \vdots \\ \beta_{1n} & \cdots & \beta_{mn} \end{pmatrix}_{n \times m}, \\
\mathbf{N_x} &= \operatorname{diag}(N_{1x}, \ldots, N_{nx})_{n \times n}, \\
\mathbf{N_y} &= \operatorname{diag}(N_{1y}, \ldots, N_{ny})_{n \times n}, \\
\mathbf{D_x} &= (D_{1x}, \ldots, D_{mx})^T, \\
\mathbf{D_y} &= (D_{1y}, \ldots, D_{my})^T, \\
\mathbf{F} &= \big((\boldsymbol{X}_1 - \boldsymbol{\mathcal{P}}_1)^T \boldsymbol{N}_1, \ldots, (\boldsymbol{X}_n - \boldsymbol{\mathcal{P}}_n)^T \boldsymbol{N}_n \big)^T.
\end{aligned}$$

3 SDM

The error term in SDM is similar to that in TDM. In SDM,

$$\begin{aligned}
e_k(\mathcal{D}) &= e_{SD,k}(\mathcal{D}) \\
&= \frac{1}{2} \frac{d_k}{d_k - \rho_k} \big[(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k)^T \boldsymbol{T}_k \big]^2 + \frac{1}{2} \big[(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k)^T \boldsymbol{N}_k \big]^2.
\end{aligned}$$

The forms of the two parts are similar. If we define $\boldsymbol{T}'_k = \sqrt{\frac{d_k}{d_k - \rho_k}} \boldsymbol{T}_k$, we can simplify the error term as:

$$e_k(\mathcal{D}) = \frac{1}{2} \big[(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k)^T \boldsymbol{T}'_k \big]^2 + \frac{1}{2} \big[(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k)^T \boldsymbol{N}_k \big]^2.$$

The minimization of $f(\mathcal{D})$ leads to a linear system as below.

$$\begin{aligned}
\text{LHS} &= \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^T \left[\begin{pmatrix} \mathbf{N_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{N_y} \end{pmatrix} \begin{pmatrix} \mathbf{N_x} & \mathbf{N_y} \\ \mathbf{N_x} & \mathbf{N_y} \end{pmatrix} + \begin{pmatrix} \mathbf{T'_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{T'_y} \end{pmatrix} \begin{pmatrix} \mathbf{T'_x} & \mathbf{T'_y} \\ \mathbf{T'_x} & \mathbf{T'_y} \end{pmatrix} \right] \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{D_x} \\ \mathbf{D_y} \end{pmatrix}, \\
\text{RHS} &= \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{pmatrix}^T \left[\begin{pmatrix} \mathbf{N_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{N_y} \end{pmatrix} \begin{pmatrix} \mathbf{F} \\ \mathbf{F} \end{pmatrix} + \begin{pmatrix} \mathbf{T'_x} & \mathbf{0} \\ \mathbf{0} & \mathbf{T'_y} \end{pmatrix} \begin{pmatrix} \mathbf{G} \\ \mathbf{G} \end{pmatrix} \right],
\end{aligned}$$

where

$$\mathbf{T}'_{\mathbf{x}} = \mathrm{diag}(\sqrt{\frac{d_1}{d_1-\rho_1}}T_{1x}, \ldots, \sqrt{\frac{d_n}{d_n-\rho_n}}T_{nx})_{n\times n},$$
$$\mathbf{T}'_{\mathbf{y}} = \mathrm{diag}(\sqrt{\frac{d_1}{d_1-\rho_1}}T_{1y}, \ldots, \sqrt{\frac{d_n}{d_n-\rho_n}}T_{ny})_{n\times n},$$
$$\mathbf{G} = \big((\boldsymbol{X}_1-\boldsymbol{\mathcal{P}}_1)^T\boldsymbol{T}'_1, \ldots, (\boldsymbol{X}_n-\boldsymbol{\mathcal{P}}_n)^T\boldsymbol{T}'_n\big)^T.$$

4 PDM

In PDM, the error term is:

$$\begin{aligned}
e_k(\mathcal{D}) &= e_{PD,k}(\mathcal{D}) \\
&= \frac{1}{2}\Big\|\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k\Big\|_2^2 \\
&= \frac{1}{2}(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k)^T(\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{X}_k) \\
&= \frac{1}{2}\boldsymbol{\mathcal{P}}_+(t_k)^T\boldsymbol{\mathcal{P}}_+(t_k) - \boldsymbol{\mathcal{P}}_+(t_k)^T\boldsymbol{X}_k + \frac{1}{2}\boldsymbol{X}_k^T\boldsymbol{X}_k.
\end{aligned}$$

The minimization of $f(\mathcal{D})$ results in two separate linear systems:

$$\mathbf{B}^T\mathbf{B}\mathbf{D}_{\mathbf{x}} = \mathbf{B}^T\mathbf{H}_{\mathbf{x}},$$
$$\mathbf{B}^T\mathbf{B}\mathbf{D}_{\mathbf{y}} = \mathbf{B}^T\mathbf{H}_{\mathbf{y}},$$

where

$$\mathbf{H}_{\mathbf{x}} = (X_{1x} - \mathcal{P}_{1x}, \ldots, X_{nx} - \mathcal{P}_{nx})^T,$$
$$\mathbf{H}_{\mathbf{y}} = (X_{1y} - \mathcal{P}_{1y}, \ldots, X_{ny} - \mathcal{P}_{ny})^T.$$

5 Internal energy terms

The internal energy terms are the integrals of the first and second derivatives of the curve:

$$F_1 = \int \|P'_D(t)\|^2 dt$$

and

$$F_2 = \int \|P''_D(t)\|^2 dt.$$

Eventually, the evaluation of the integration of the form $\int \beta_{ik}\beta_{jk}dt$ is needed. Here, I focus on the special case of cardinal spline basis. It seems that there would be no explicit/analytical form of these integrals.

6 Appendix

Some useful relations are listed here, in case I forget the derivation process.

6.1 matrix manipulation

$$\sum_j^m \sum_k^n \beta_{ik}\beta_{jk}N_kM_kD_{jx}$$
$$= \begin{pmatrix}\beta_{i1}N_1M_1 & \dots & \beta_{in}N_nM_n\end{pmatrix} \mathbf{B} \begin{pmatrix} D_{1x} \\ \vdots \\ D_{mx} \end{pmatrix}$$

$$\begin{pmatrix} \beta_{11}N_{1x}N_{1y} & \dots & \beta_{1n}N_{nx}N_{ny} \\ \vdots & \ddots & \vdots \\ \beta_{m1}N_{1x}N_{1y} & \dots & \beta_{mn}N_{nx}N_{ny} \end{pmatrix}$$
$$= \mathbf{B}^T \cdot \mathrm{diag}(N_{1x}, \dots, N_{nx}) \cdot \mathrm{diag}(N_{1y}, \dots, N_{ny})$$

6.2 manipulation of cardinal splines

For cardinal splines $\$_{j,k,\mathbb{Z}}$, there is only one "unique" spline for a given order k; all other cardinal splines of order k can be generated by translating the one that spans $[0, k+1)$. I denotes these "unique" cardinal splines as Q_k of order k. I list the cardinal splines up to order 4.

$$Q_1 = \begin{cases} 1 & 0 \le x < 1 \\ 0 & otherwise \end{cases},$$
$$Q_2 = \begin{cases} x & 0 \le x < 1 \\ -(x-1)+1 & 1 \le x < 2 \\ 0 & otherwise \end{cases},$$
$$Q_3 = \begin{cases} \frac{1}{2}x^2 & 0 \le x < 1 \\ -(x-1)^2 + (x-1) + \frac{1}{2} & 1 \le x < 2 \\ \frac{1}{2}(x-2)^2 - (x-2) + \frac{1}{2} & 2 \le x < 3 \\ 0 & otherwise \end{cases},$$
$$Q_4 = \begin{cases} \frac{1}{6}x^3 & 0 \le x < 1 \\ -\frac{1}{2}(x-1)^3 + \frac{1}{2}(x-1)^2 + \frac{1}{2}(x-1) + \frac{1}{6} & 1 \le x < 2 \\ \frac{1}{2}(x-2)^3 - (x-2)^2 + \frac{2}{3} & 2 \le x < 3 \\ -\frac{1}{6}(x-3)^3 + \frac{1}{2}(x-3)^2 - \frac{1}{2}(x-3) + \frac{1}{6} & 3 \le x < 4 \\ 0 & otherwise \end{cases}.$$

The derivative of the Q_k is:

$$Q'_{i,k} = Q_{i,k-1} - Q_{i+1,k-1},$$

where

$$Q_{i,k} = Q_k(x-i),$$

e.g., $Q_{i,k}$ is a translate of Q_k.

Suppose that a order k b-spline curve with cardinal basis Q_j is given: $\boldsymbol{\mathcal{P}} = \sum_{j=1}^m \boldsymbol{P}_j Q_j$. The first and second derivatives of $\boldsymbol{\mathcal{P}}$ with respect to x are:

$$\boldsymbol{\mathcal{P}}' = \sum_{j=1}^{m+1} [Q_{j,k-1}(\boldsymbol{P}_j - \boldsymbol{P}_{j-1})],$$
$$\boldsymbol{\mathcal{P}}'' = \sum_{j=1}^{m+2} [Q_{j,k-2}(\boldsymbol{P}_j - 2\boldsymbol{P}_{j-1} + \boldsymbol{P}_{j-2})].$$

For open curves,

$$\boldsymbol{P}_j = 0, j \notin [1, m],$$

while for closed curves,

$$\boldsymbol{P}_j = \boldsymbol{P}_{j \bmod m}, j \notin [1, m].$$

The derivatives of $\boldsymbol{\mathcal{P}}'_+$ and $\boldsymbol{\mathcal{P}}''_+$ with respect to D_{ix} and D_{iy} are:

$$\frac{\partial \boldsymbol{\mathcal{P}}'_+}{\partial D_{ix}} = [Q_{i,k-1} - Q_{i+1,k-1}, 0]^T$$
$$\frac{\partial \boldsymbol{\mathcal{P}}'_+}{\partial D_{iy}} = [0, Q_{i,k-1} - Q_{i+1,k-1}]^T$$
$$\frac{\partial \boldsymbol{\mathcal{P}}''_+}{\partial D_{ix}} = [Q_{i,k-2} - 2Q_{i+1,k-2} + Q_{i+2,k-2}, 0]^T$$
$$\frac{\partial \boldsymbol{\mathcal{P}}''_+}{\partial D_{iy}} = [0, Q_{i,k-2} - 2Q_{i+1,k-2} + Q_{i+2,k-2}]^T.$$

For open curves,

$$Q_j = 0, j \notin [1, m],$$

while for closed curves,

$$Q_j = Q_{j \bmod m}, j \notin [1, m].$$

Two regularization terms (F_1 and F_2) are added in cost function f to control the smoothness of the b-spline curves:

$$F_1 = \frac{1}{2} \int_{\boldsymbol{\mathcal{P}}} \|\boldsymbol{\mathcal{P}}'\|^2 dt,$$
$$F_2 = \frac{1}{2} \int_{\boldsymbol{\mathcal{P}}} \|\boldsymbol{\mathcal{P}}''\|^2 dt.$$

Now we calculate the partial derivative of F_1 with respect to D_{ix} and D_{iy}:

$$\begin{aligned}
\frac{\partial F_1}{\partial D_{ix}} &= \frac{\partial}{\partial D_{ix}} \frac{1}{2} \int ({P'_{x+}}^2 + {P'_{y+}}^2) dt \\
&= \frac{1}{2} \int \frac{\partial}{\partial D_{ix}} ({P'_{x+}}^2 + {P'_{y+}}^2) dt \\
&= \int (P'_{x+} \cdot \frac{\partial P'_{x+}}{\partial D_{ix}} + P'_{y+} \cdot \frac{\partial P'_{y+}}{\partial D_{ix}}) dt,
\end{aligned}$$

where

$$P_{x+} = \sum_{j=1}^{m} (P_{jx} + D_{jx}) Q_{j,k},$$

$$P_{y+} = \sum_{j=1}^{m} (P_{jy} + D_{jy}) Q_{j,k},$$

$$\frac{\partial P'_{x+}}{\partial D_{ix}} = Q_{i,k-1} - Q_{i+1,k-1}$$

and

$$\frac{\partial P'_{y+}}{\partial D_{ix}} = 0.$$

After simplify the above equation, we have

$$\begin{aligned}
\frac{\partial F_1}{\partial D_{ix}} &= \int (P'_{x+} \cdot \frac{\partial P'_{x+}}{\partial D_{ix}} + P'_{y+} \cdot \frac{\partial P'_{y+}}{\partial D_{ix}}) dt \\
&= \int (P'_{x+} \cdot \frac{\partial P'_{x+}}{\partial D_{ix}}) dt \\
&= \sum_{j=1}^{m+1} \left\{ \Big[(P_{jx} + D_{jx}) - (P_{j-1,x} + D_{j-1,x}) \Big] \cdot \int \Big[Q_{j,k-1} (Q_{i,k-1} - Q_{i+1,k-1}) \Big] dt \right\}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\frac{\partial F_1}{\partial D_{iy}} &= \sum_{j=1}^{m+1} \left\{ \Big[(P_{jy} + D_{jy}) - (P_{j-1,y} + D_{j-1,y}) \Big] \cdot \int \Big[Q_{j,k-1} (Q_{i,k-1} - Q_{i+1,k-1}) \Big] dt \right\}, \\
\frac{\partial F_2}{\partial D_{ix}} &= \sum_{j=1}^{m+2} \left\{ \Big[(P_{jx} + D_{jx}) - 2(P_{j-1,x} + D_{j-1,x}) + (P_{j-2,x} + D_{j-2,x}) \Big] \right. \\
&\qquad \left. \cdot \int \Big[Q_{j,k-2} (Q_{i,k-2} - 2Q_{i+1,k-2} + Q_{i+2,k-2}) \Big] dt \right\}, \\
\frac{\partial F_2}{\partial D_{iy}} &= \sum_{j=1}^{m+2} \left\{ \Big[(P_{jy} + D_{jy}) - 2(P_{j-1,y} + D_{j-1,y}) + (P_{j-2,y} + D_{j-2,y}) \Big] \right. \\
&\qquad \left. \cdot \int \Big[Q_{j,k-2} (Q_{i,k-2} - 2Q_{i+1,k-2} + Q_{i+2,k-2}) \Big] dt \right\}.
\end{aligned}$$