Example 1: Least Squares (1)

$A(m\times n), \hat{u}(n\times 1), b(m\times 1), m>n$

Minimize $E=\|e\|^2=\|Au-b\|^2$

$\rightarrow A^T A\hat{u}=A^T b$ (normal equation for best $\hat{u}$)

[normal equations]

$A^T e=0$: e as the projection of b across onto the null space of A^T (dimension $m-n$)

$A^T(b-A\hat{u})=0$: $A\hat{u}$ as the projection of b down onto the column space of A (dimension n)

[Saddle Point/ Kuhn-Tucker(KKT)/ Primal-Dual]

$$\left.\begin{aligned} A^T(b-A\hat{u})&=0 \\ A^T e&=0 \end{aligned}\right\} \rightarrow \left.\begin{aligned} m:&\quad e+A\hat{u}=b \\ n:&\quad A^T e\quad\;\;=0 \end{aligned}\right\} \rightarrow \underbrace{\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix}}_{\substack{\text{saddle point matrix} \\ \text{KKT matrix}}} \begin{bmatrix} e \\ \hat{u} \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

Least Squares (2)

- Saddle point

$$S = \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \rightarrow \begin{bmatrix} I & 0 \\ -A^T & I \end{bmatrix} \begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} = \begin{bmatrix} I & A \\ 0 & -A^T A \end{bmatrix} \rightarrow \text{indefinite}$$

- Kuhn-Tucker (KKT)
  - In continuous problems, Euler-Lagrange equations
- Primal-Dual

$$\begin{cases} \text{primal} : \text{Minimize } \dfrac{1}{2}\|Au - b\|^2 \rightarrow \text{solution } \hat{u} \\ \text{dual} : \text{Minimize } \dfrac{1}{2}\|e - b\|^2 \ \text{ with } A^T e = 0 \rightarrow \text{Lagrange multiplier } u \end{cases}$$

Weighted Least Squares

$$\text{Minimize } E = \|e\|^2 = \|WAu - Wb\|^2 \rightarrow (WA)^T (WA)\hat{u}_W = (WA)^T (Wb)$$
$$\xrightarrow{C=W^TW} A^T CA\hat{u}_W = A^T Cb \quad \text{(normal equation)}$$

[normal equations]

$A^T Ce = 0$: e as the projection of b across onto the null space of $A^T C$ (dimension $m-n$)

$A^T C(b - A\hat{u}_W) = 0$: $A\hat{u}_W$ as the projection of b down onto the column space of A (dimension n)

[Saddle Point/ Kuhn-Tucker(KKT)/ Primal-Dual]

$$\left.\begin{array}{l} A^T C(b - A\hat{u}_W) = 0 \\ A^T Ce = 0 \end{array}\right\} \rightarrow \left\{\begin{array}{lr} m: \quad e + A\hat{u}_W & = b \\ n: \quad A^T Ce & = 0 \end{array}\right\} \xrightarrow{w=Ce} \left\{\begin{array}{lr} C^{-1}w + A\hat{u}_W & = b \\ A^T w & = 0 \end{array}\right.$$

$$\rightarrow \begin{bmatrix} C^{-1} & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} w \\ \hat{u}_W \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} C^{-1} & A \\ 0 & -A^T CA \end{bmatrix} \begin{bmatrix} w \\ \hat{u}_W \end{bmatrix} = \begin{bmatrix} b \\ -A^T Cb \end{bmatrix}$$

Duality

$$\begin{cases} \text{Problem 1 (project down to } A\hat{u}\text{)}: \text{Minimize } \|b - Au\|^2 \\ \text{Problem 2 (project across to } e\text{)}: \text{Minimize } \|b - w\|^2 \text{ with } A^T w = 0 \end{cases}$$

$$Au \perp w$$

$$\|b - Au\|^2 + \|b - w\|^2 = \|b\|^2 + \|b - Au - w\|^2 \xrightarrow{u=\hat{u},w=e \rightarrow b=A\hat{u}+e}$$

$$\|b - A\hat{u}\|^2 + \|b - e\|^2 = \|b\|^2$$

Optimality: $\left(b - A\hat{u}\right)$ in Problem $1 = e$ in Problem 2

Duality: $\|b - Au\|^2 + \|b - w\|^2 - \|b\|^2 \geq 0 \rightarrow \underbrace{\|b - w\|^2 - \|b\|^2}_{\min\ @\ w=e} \geq \underbrace{-\|b - Au\|^2}_{\max\ @\ u=\hat{u}}$

Example 2: Minimizing with Constraints

- Elimination / Idea of Lagrange

Internal energy in the springs: $E(w) = E_1(w_1) + E_2(w_2)$

Balance of internal/external forces: $w_1 - w_2 = f$

$\rightarrow$ Constrained optimization problem: Minimize $E(w)$ subject to $w_1 - w_2 = f$

Lagrange function: $L(w_1, w_2, u) = E_1(w_1) + E_2(w_2) - u(w_1 - w_2 - f)$

$$(\text{Kuhn-Tucker optimality equations}) \begin{cases} \dfrac{\partial L}{\partial w_1} = \dfrac{\partial E_1}{\partial w_1} - u = 0 \\ \dfrac{\partial L}{\partial w_2} = \dfrac{\partial E_2}{\partial w_2} + u = 0 \\ \dfrac{\partial L}{\partial u} = -(w_1 - w_2 - f) = 0 \end{cases}$$

Lagrange multiplier u : sensitivity of the answer to a change in the constraint

- displacement, selling price

Linear Case

$$\text{linear spring: } E = \frac{1}{2} c e^2 \xrightarrow{w=ce} E = \frac{1}{2} \frac{w^2}{c}$$

$$\text{Minimize } E(w) = \frac{1}{2} \frac{w_1^2}{c_1} + \frac{1}{2} \frac{w_2^2}{c_2} \text{ subject to } w_1 - w_2 = f$$

- Geometry
  - Line is tangent to the ellipse
- Algebra
  - Energy gradient: Ku = f
  - Spring forces: $E_{\min} = \frac{1}{2} f^T K^{-1} f$
  - Lagrange multiplier = sensitivity

Fundamental Problem

Minimize the total energy $E(w) = \frac{1}{2} w^T C^{-1} w$ in the m springs

subject to n balance of forces $A^T w = f$

Lagrange function: $L = \frac{1}{2} w^T C^{-1} w - u^T \left(A^T w - f \right)$

Optimality conditions: $\begin{cases} \partial L / \partial w = C^{-1} w - Au = 0 \rightarrow w = CAu \\ \partial L / \partial u = -A^T w + f = 0 \rightarrow A^T w = f \end{cases}$

$\rightarrow A^T CAu = f$

Least squares problems: $e = b - Au \rightarrow E(w) = \frac{1}{2} w^T C^{-1} w - b^T w$

$$L = \frac{1}{2} w^T C^{-1} w - b^T w + u^T \left(A^T w - f \right)$$

$$\left.\begin{cases} \partial L / \partial w = C^{-1} w + Au - b = 0 \\ \partial L / \partial u = A^T w - f = 0 \end{cases}\right\} \rightarrow \begin{bmatrix} C^{-1} & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} b \\ f \end{bmatrix}$$

$$\rightarrow A^T w = A^T C \left(b - Au \right) = f \rightarrow A^T CAu = A^T Cb - f$$

8.2 Regularized Least Squares

[ordinary least squares]

Minimize $\|Au-b\|^2 \rightarrow A^T A\hat{u} = A^T b \rightarrow$ can be "ill-posed"$(\alpha$=0, A: highly ill-conditioned$)$

[weighted least squares]

Minimize $(b-Au)^T C(b-Au) \rightarrow A^T CA\hat{u} = A^T Cb$

[regularized least squares: special case of weighted least squares]

$$\text{Minimize } \|Au-b\|^2 + \alpha\|Bu-d\|^2 \rightarrow \left(A^T A + \underbrace{\alpha B^T B}_{\text{regularize } A^T A} \right)\hat{u} = A^T b + \alpha B^T d$$

$$\leftrightarrow \begin{bmatrix} A^T & B^T \end{bmatrix} \underbrace{\begin{bmatrix} I & 0 \\ 0 & \alpha I \end{bmatrix}}_{C} \begin{bmatrix} A \\ B \end{bmatrix} \hat{u} = \begin{bmatrix} A^T & B^T \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \alpha I \end{bmatrix} \begin{bmatrix} B \\ d \end{bmatrix}$$

$\alpha \uparrow \rightarrow \|\hat{u}\| \downarrow, \|Au-b\| \uparrow \rightarrow \alpha : \|Au-b\| \approx$ (expected noise)

[constrained least squares]

Minimize $\|Au-b\|^2$ subject to $Bu = d$: $\alpha \rightarrow \infty, \|B\hat{u}_\alpha - d\|^2 \rightarrow 0$

Minimize $\|Au-b\|^2$ subject to $Bu=d$

$[1]$ large penalty: $\hat{u}_\infty$ minimizes $\|Au-b\|^2$ among all minimizers of $\|Bu-d\|^2$

$[2]$ Lagrange multuplier: $L(u,w)=\frac{1}{2}\|Au-b\|^2+w(Bu-d)$

$$\rightarrow \frac{\partial L}{\partial u}=\frac{\partial L}{\partial w}=0 \rightarrow \begin{bmatrix} A^T A & B^T \\ B & 0 \end{bmatrix}\begin{bmatrix} u \\ w \end{bmatrix}=\begin{bmatrix} A^T b \\ d \end{bmatrix}$$

$[3]$ nullspace method: solve $Bu=d$ for $u=u_n+u_r \rightarrow \begin{Bmatrix} Bu_n=0 \\ Bu_r=d \end{Bmatrix} \rightarrow u=Q_n z+u_r$

minimizes $\|A(u_n+u_r)-b\|^2=\|AQ_n z-(b-Au_r)\|^2$

solve $(n-p)$ normal equations $AQ_n z=b-Au_r \rightarrow {Q_n}^T A^T AQ_n\hat{z}={Q_n}^T A^T(b-Au_r)$

$$qr(B') \rightarrow B^T=[Q_r \quad Q_n]\begin{bmatrix} R \\ 0 \end{bmatrix} \rightarrow BQ_n=[R^T \quad 0]\begin{bmatrix} {Q_r}^T \\ {Q_n}^T \end{bmatrix}Q_n==[R^T \quad 0]\begin{bmatrix} 0 \\ I \end{bmatrix}=0$$

$$Bu_r=[R^T \quad 0]\begin{bmatrix} {Q_r}^T \\ {Q_n}^T \end{bmatrix}u_r=d \rightarrow u_r=Q_r\left(R^{-1}\right)^T d=B^+ d$$

Minimize $\|Au\|^2 = u_1^2 + u_2^2$ subject to $Bu = u_1 - u_2 = 8$

$$\leftrightarrow A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ b = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ B = \begin{bmatrix} 1 & -1 \end{bmatrix},\ d = \begin{bmatrix} 8 \end{bmatrix}$$

$[1]$ large penalty: minimize $u_1^2 + u_2^2 + \alpha(u_1 - u_2 - 8)^2$ and let $\alpha \to \infty$

$[2]$ Lagrange multuplier: find a saddle point of $L = \frac{1}{2}\left(u_1^2 + u_2^2\right) + w(u_1 - u_2 - 8)$

$[3]$ nullspace method: solve $Bu = d$ and look for the shortest solution

Pseudoinverse

$u^{+} = A^{+}b$ is the shortest vector that solves $A^{T}Au^{+} = A^{T}b$

A is square and invertible $\rightarrow u = A^{-1}b : A^{+} = A^{-1}$

A is rectangular and has independent columns $\rightarrow \hat{u} = \left(A^{T}A\right)^{-1}A^{T}b : A^{+} = \left(A^{T}A\right)^{-1}A^{T}$

$$A = U\Sigma V^{T} = \begin{bmatrix} \underbrace{U_{col}}_{r \text{ columns}} & U_{null} \end{bmatrix} \begin{bmatrix} \Sigma_{pos} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{row} & V_{null} \end{bmatrix}^{T}$$

$$A^{+} = \left(V_{row}\right)\left(\Sigma_{pos}\right)^{-1}\left(U_{col}\right)^{T}$$

8.3 Calculus of Variations

minimize $P(u) = \frac{1}{2} u^T K u - u^T f \rightarrow P' = Ku - f = 0$

minimize $P(u) = \int \left(\frac{1}{2} c \left(u'(x) \right)^2 - f(x) u(x) \right) dx \rightarrow \frac{\delta P}{\delta u} = 0$

perturb $u(x)$ by a test function $v(x)$

compare $P(u)$ with $P(u+v) \rightarrow \frac{\delta P}{\delta u} = 0$ for every admissible v

Euler-Lagrange equation $\rightarrow \begin{cases} \text{weak form for every } v\text{: } \int c u' v' dx = \int f v dx \\ \text{strong form for every point: } -\left(c u' \right)' = f \end{cases}$

1D Problem

$$P(u) = \int_0^1 F(u, u')dx \quad \text{with } u(0) = a \text{ and } u(1) = b$$

$$F(u + v, u' + v') = F(u, u') + v\frac{\partial F}{\partial u} + v'\frac{\partial F}{\partial u'} + \cdots$$

$$P(u + v) = P(u) + \int_0^1 \left(v\frac{\partial F}{\partial u} + v'\frac{\partial F}{\partial u'} \right) dx + \cdots$$

$$\langle \text{weak form} \rangle \quad \frac{\delta P}{\delta u} = \int_0^1 \left(v\frac{\partial F}{\partial u} + v'\frac{\partial F}{\partial u'} \right) dx = 0 \quad \text{for every } v$$

$$\left[v\frac{\partial F}{\partial u'} \right]_0^1 + \int_0^1 v\left(\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'} \right) \right) dx = 0$$

$$\langle \text{strong form} \rangle \quad \frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'} \right) = 0: \text{ Euler-Lagrange equation for } u$$

constraints on u: Lagrange multipliers and saddle points as L

$$P = \int_0^1 F\left(x,u,u'\right)dx \quad \text{with } u\left(0\right) = a \text{ and } u\left(1\right) = b$$

$$P\left(\varepsilon\right) = \int_0^1 F\left(x,u+\varepsilon v,u'+\varepsilon v'\right)dx$$

$$\frac{dP}{d\varepsilon} = \int_0^1 \frac{dF\left(x,u+\varepsilon v,u'+\varepsilon v'\right)}{d\varepsilon}dx = \int_0^1 \left[\frac{\partial F}{\partial x}\frac{dx}{d\varepsilon} + \frac{\partial F}{\partial\left(u+\varepsilon v\right)}\frac{d\left(u+\varepsilon v\right)}{d\varepsilon} + \frac{\partial F}{\partial\left(u'+\varepsilon v'\right)}\frac{d\left(u'+\varepsilon v'\right)}{d\varepsilon}\right]dx$$

$$\left.\frac{dP}{d\varepsilon}\right|_{\varepsilon=0} = \int_0^1 \left(\frac{\partial F}{\partial u}v + \frac{\partial F}{\partial u'}v'\right)dx = 0 \quad \text{at extremum}$$

$$\frac{\partial F}{\partial u} - \frac{d}{dx}\left(\frac{\partial F}{\partial u'}\right) = 0 \rightarrow \frac{1}{u'}\left[\frac{d}{dx}\left(F - \frac{\partial F}{\partial u'}\frac{du}{dx}\right) - \frac{\partial F}{\partial x}\right]$$

$$\begin{cases} F = F\left(u,u'\right) \rightarrow \dfrac{\partial F}{\partial x} = 0: \quad F - \dfrac{\partial F}{\partial u'}\dfrac{du}{dx} = c \\ F = F\left(x,u'\right) \rightarrow \dfrac{\partial F}{\partial u} = 0: \quad \dfrac{\partial F}{\partial u'} = c \end{cases}$$

Example

Find the shortest path $u(x)$ between $(0,a)$ and $(1,b)$: $u(0)=a$ and $u(1)=b$

When the constraint is $\int u(x)\,dx = A$, find the shortest curve that has area A below it

2D Problem

$$P(u)=\iint_S\left[\frac{c}{2}\left(\frac{\partial u}{\partial x}\right)^2+\frac{c}{2}\left(\frac{\partial u}{\partial y}\right)^2-f(x,y)u(x,y)\right]dxdy:\ \text{potential energy over a plane region } S$$

$$\langle\text{weak form}\rangle\quad \frac{\delta P}{\delta u}=\iint_S\left[c\left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial v}{\partial x}\right)+c\left(\frac{\partial u}{\partial y}\right)\left(\frac{\partial v}{\partial y}\right)-fv\right]dxdy=0$$

$$\rightarrow\iint_S\left[-\frac{\partial}{\partial x}\left(c\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial y}\left(c\frac{\partial u}{\partial y}\right)-f\right]vdxdy=0$$

$$\langle\text{strong form}\rangle\quad -\frac{\partial}{\partial x}\left(c\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial y}\left(c\frac{\partial u}{\partial y}\right)=f:\ \text{Euler-Lagrange equation}$$

$$\rightarrow A^TCAu=-\nabla\bullet c\nabla u=f\ \text{ with } Au=\nabla u$$

$$P(u)=\frac{1}{2}\int c(u')^2dx-\int fudx=\frac{1}{2}a(u,u)-l(u)$$

weak form: $a(u,v)=l(v)$ for all admissible v

Galerkin method: $a(U,V)=l(V)$ for all test functions V

Minimal Surface

$$E(u) = \iint_S \sqrt{1+\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2}\,dxdy: \text{ surface area}$$

$$F = 1+\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2, G = 2\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+2\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}+O\left(v^2\right)$$

$$\sqrt{F+G} = \sqrt{F}+\frac{G}{2\sqrt{F}}+\cdots$$

$$E\left(u+v\right) = E(u)+\iint_S \frac{1}{\sqrt{F}}\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dxdy+\cdots$$

$$\left\langle\text{weak form}\right\rangle \quad \frac{\delta E}{\delta u} = \iint_S \frac{1}{\sqrt{F}}\left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial v}{\partial y}\right)dxdy = 0 \quad \text{for every } v$$

$$\left\langle\text{strong form}\right\rangle \quad -\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{F}}\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial y}\left(\frac{1}{\sqrt{F}}\frac{\partial u}{\partial y}\right) = 0: \text{ Euler-Lagrange equation}$$

Nonlinear Equations

Minimize $E = \iint F dxdy$ where $F = F\left(x, y, u\left(= D_0 u\right), D_1 u\left(= \frac{du}{dx}\right), D_2 u\left(= \frac{du}{dy}\right), \ldots \right)$

$$F\left(u+v\right) = F\left(u\right) + F'\left(u\right)v + O\left(v^2\right) = F\left(u\right) + \sum \frac{\partial F}{\partial D_i u} D_i v + \cdots = F\left(x, y, D_0 u + D_0 v, D_1 u + D_1 v, \ldots\right)$$

$\langle$weak form$\rangle$ $\frac{\delta E}{\delta u} = \iint_S \left(\sum \frac{\partial F}{\partial D_i u} \right) \left(D_i v\right) dxdy = 0$ for every v

$\langle$strong form$\rangle$ $\sum D_i^T \left(\frac{\partial F}{\partial D_i u} \right) = \sum D_i^T w_i = 0$: Euler-Lagrange equation