1.4 Minimum Principles

Notes

1E If A is positive definite, the quadratic $P(x) = \frac{1}{2}x^TAx - x^Tb$ is minimized at the point where $Ax=b$. The minimum value is

$$P(A^{-1}b) = -\frac{1}{2}b^TA^{-1}b$$

Proof: Suppose x is the solution to $Ax=b$ and let y be any point.

$$\begin{aligned}
P(y) - P(x) &= \tfrac{1}{2}y^TAy - y^Tb - \tfrac{1}{2}x^TAx + x^Tb \\
&= \tfrac{1}{2}y^TAy - y^TAx - \tfrac{1}{2}x^TAx + x^TAx \\
&= \tfrac{1}{2}y^TAy - y^TAx + \tfrac{1}{2}x^TAx \\
&= \tfrac{1}{2}y^TAy - \tfrac{1}{2}y^TAx - \tfrac{1}{2}y^TAx + \tfrac{1}{2}x^TAx \\
&= \tfrac{1}{2}(y^TAy - y^TAx - y^TAx + x^TAx) \\
&= \tfrac{1}{2}(y^TAy - x^TAy - y^TAx + x^TAx) \\
&= \tfrac{1}{2}((y^T - x^T)Ay - (y^T - x^T)Ax) \\
&= \tfrac{1}{2}((y-x)^TAy + (y-x)^TA(-x)) \\
&= \tfrac{1}{2}(y-x)^TA(y-x)
\end{aligned}$$

Use: $y^TAx = y^TA^Tx = x^TAy$

$P(y) - P(x) = \frac{1}{2}(y-x)^TA(y-x) \geq 0$ with equality iff $y - x = 0$ since A is positive definite. Conclude that $x = A^{-1}b$ is the unique minimizer of P and

$$\begin{aligned}
P(x_{min}) = P(A^{-1}b) &= \tfrac{1}{2}(A^{-1}b)^TAA^{-1}b - (A^{-1}b)^Tb \\
&= \tfrac{1}{2}b^T(A^{-1})^Tb - b^T(A^{-1})^Tb \\
&= \tfrac{1}{2}b^TA^{-1}b - b^TA^{-1}b \\
&= -\tfrac{1}{2}b^TA^{-1}b
\end{aligned}$$

Use: $I = AA^{-1} = A^TA^{-1} \rightarrow (A^T)^{-1} = A^{-1}$

Alternatively,

$$\frac{1}{2}(x - A^{-1}b)^TA(x - A^{-1}b) - \frac{1}{2}b^TA^{-1}b$$

$-\frac{1}{2}b^TA^{-1}b$ constant wrt x. Writing $P(x)$ in this way shows $P(x)$ minimized when $x = A^{-1}b$ and $P_{min} = -\frac{1}{2}b^TA^{-1}b$.

$$= \tfrac{1}{2}x^TAx - \tfrac{1}{2}b^T(A^{-1})^TAx - \tfrac{1}{2}x^TAA^{-1}b + \tfrac{1}{2}b^T(A^{-1})^TAA^{-1}b - \tfrac{1}{2}b^TA^{-1}b$$

$$= \tfrac{1}{2}x^TAx - \tfrac{1}{2}b^T(A^{-1})^Tb - \tfrac{1}{2}x^Tb + 0 = \tfrac{1}{2}x^TAx - \tfrac{1}{2}x^Tb - \tfrac{1}{2}x^Tb = \tfrac{1}{2}x^TAx - x^Tb = P(x)$$

Consider now the case that A is not assumed to be symmetric positive definite. The coefficient matrix for a physical problem often gets assembled as A^TCA. The matrices A^TA and A^TCA are always symmetric (when C is). We want to know when they are positive definite.

1F (i) If A has linearly independent columns — it can be square or rectangular — then the product A^TA is positive definite.

(ii) If C is symmetric positive definite, so is the triple product A^TCA.

Note that if A is $m \times n$, the columns of A can only be independent if $n \le m$. If $n > m$ we have no hope of independence since the number of independent rows is the same as the number of independent columns by the fundamental theorem of linear algebra.

Proof (1F): Suppose the columns of A are linearly independent.

$x^TA^TAx = (Ax)^T(Ax) = \|Ax\|_2^2 \geq 0$ with equality iff $Ax = 0$. Since the columns of A are independent, $Ax = 0$ iff $x = 0$. Thus x^TA^TAx is positive except when $x = 0$. Conclude A^TA is positive definite.

Suppose C is positive definite and A has linearly independent columns.

$x^TA^TCAx = (Ax)^TC(Ax) \geq 0$ with equality iff $Ax = 0$ iff $x = 0$.

Therefore A^TCA is positive definite.

Least Squares Solution of $Ax = b$

Suppose A is $m \times n$ with $m > n$. The problem $Ax = b$ is an overdetermined system of m equations in n unknowns. The vectors b that can be solved for form the n-dimensional subspace of m-dimensional space. This subspace is called the column space of A. $Ax = b$ has a solution iff b is in the column space of A. With $n < m$ that is unlikely.

In general there will be an error $e = b - Ax$. We emphasize that the components e_i of e are the 'vertical' distances between each component of b and Ax, not the shortest distances.

To minimize error is equivalent to minimizing $\|Ax - b\|^2 = (Ax - b)^T(Ax - b)$ if we measure a vector in the most common way.

The question is which x will minimize $\|Ax - b\|^2$. The answer is given in 1G, which we prove using 1E.

1G The x that minimizes $\|Ax-b\|^2$ is the solution to the <u>normal equations</u>:

$$A^TAx = A^Tb$$

This vector $x = (A^TA)^{-1}A^Tb$ is the least squares solution to $Ax=b$.

Proof: For this section it is assumed that the columns of A are independent. Then A^TA is positive definite by 1F. It follows that A^TA is invertible since

$A^TAx = 0 \rightarrow x^TA^TAx = 0 \rightarrow x = 0$ shows that if x is in the null space of A, then $x = 0$. In other words, the null space of A^TA is trivial.

The error $e = b - Ax$ is minimized when $\|Ax-b\|^2$ is minimized.

$$\|Ax-b\|^2 = (Ax-b)^T(Ax-b) = (x^TA^T - b^T)(Ax-b) = x^TA^TAx - x^TA^Tb - b^TAx + b^Tb$$

But $\|Ax-b\|^2$ is minimized when $\frac{1}{2}\|Ax-b\|^2$ is minimized and $x^TA^Tb = b^TAx$ since both are scalars so $x^TA^Tb = (x^TA^Tb)^T = b^TAx$. Finally note that since b^Tb is constant wrt x, it does not affect our choice of x in the minimization and should therefore be ignored in the minimization.

Thus we want to find the x that minimizes $\frac{1}{2}x^TA^TAx - x^TA^Tb$. Recall from 1E that the solution x of $Ax=b$ minimizes $P(x) = \frac{1}{2}x^TAx - x^Tb$. Replace the matrix A by the matrix A^TA and the vector b by the vector A^Tb. Then 1E applies here since A^TA is positive definite. This tells us that the x that minimizes $\frac{1}{2}x^TA^TAx - x^TA^Tb$ is the solution to $A^TAx = A^Tb$.

$$\therefore\ x = (A^TA)^{-1}A^Tb \text{ minimizes the error } e = b - Ax$$

We could call the error minimizing point $p = (A^TA)^{-1}A^Tb$ for 'projection' since Ap is the closest point to b in the column space of A. Note

$$A^Te = A^T(b - Ap) = A^Tb - A^TAp = 0$$

The inner (dot) product of e with each column of A is zero by this calculation — the error is orthogonal to the column space of A when we choose $x = p$, the projection of b onto the column space of A.

The takeaway from this section is simple:

<u>If $Ax=b$ has no solution, multiply by A^T and solve $A^TAx = A^Tb$</u>

Linear Regression

1H Given measurements $b_1, \dots, b_m$ at times $t_1, \dots, t_m$, the line $y = \alpha t + \beta$ which minimizes error $\|e\|^2$ is determined by

$$A^T A x = A^T b \quad \text{or} \quad \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

The best line is $y = \bar{b} + \alpha (t - \bar{t})$ with $\alpha = \dfrac{\sum (t_i - \bar{t}) b_i}{\sum (t_i - \bar{t})^2}$

Proof: The previous section explains how to set up a system that will minimize the error in describing the relationship between the t_i and b_i using a line $y = Ct + D$. First write out the system we want to solve

$$\begin{matrix} \beta + \alpha t_1 = b_1 \\ \beta + \alpha t_2 = b_2 \\ \vdots \\ \beta + \alpha t_m = b_m \end{matrix} \longrightarrow \underset{(A)}{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}} \underset{(x)}{\begin{bmatrix} \beta \\ \alpha \end{bmatrix}} = \underset{(b)}{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}} \longrightarrow Ax = b$$

If $Ax = b$ has a solution this is fine – it means all the b_i lie along a line and error $e = 0$. But in the case we mainly want to consider, the measurements will not be perfectly linear and $Ax = b$ has no solution:

$$\text{solve } A^T A x = A^T b \text{ for } x$$

Unless $t_1 = t_2 = \dots = t_m$ the columns of A are independent, $A^T A$ is positive definite and thus a unique solution $x = (A^T A)^{-1} A^T b$ can be found.*

$$A^T A x = \begin{bmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = A^T b$$

$$A^T A x = \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum b_i t_i \end{bmatrix} = A^T b$$

This shows that to find the α, β such that the error in fitting $y = \alpha t + \beta$ to b is minimized by solving this 2×2 system. Use either the formula for the inverse of a 2×2 matrix or use Gaussian elimination. We use the second approach next to find $x = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$.

* If the columns of A were dependent we would need to use a pseudoinverse.

$$\left[\begin{array}{cc|c} m & \Sigma t_i & \Sigma b_i \\ \Sigma t_i & \Sigma t_i^2 & \Sigma t_i b_i \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & \bar{t} & \bar{b} \\ \bar{t} & \frac{1}{m}\Sigma t_i^2 & \frac{1}{m}\Sigma t_i b_i \end{array}\right]$$

$$\sim \left[\begin{array}{cc|c} 1 & \bar{t} & \bar{b} \\ 0 & \frac{1}{m}\Sigma t_i^2 - \bar{t}^2 & \frac{1}{m}\Sigma t_i b_i - \bar{b}\bar{t} \end{array}\right]$$

$$\sim \left[\begin{array}{cc|c} 1 & \bar{t} & \bar{b} \\ 0 & 1 & \alpha \end{array}\right] \quad (\alpha \text{ simplified below})$$

$$\sim \left[\begin{array}{cc|c} 1 & 0 & \beta \\ 0 & 1 & \alpha \end{array}\right] \quad (\beta = \bar{b} - \alpha\bar{t} \text{ simplified below})$$

$$\alpha = \frac{\frac{1}{m}\Sigma t_i b_i - \bar{b}\bar{t}}{\frac{1}{m}\Sigma t_i^2 - \bar{t}^2} = \frac{\Sigma t_i b_i - \frac{1}{m}\Sigma b_i \Sigma t_i}{\Sigma t_i^2 - \frac{1}{m}\Sigma t_i \Sigma t_i} = \frac{\Sigma t_i b_i - \Sigma \bar{t} b_i}{\Sigma t_i^2 - \Sigma \bar{t} t_i} = \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i^2 - \bar{t} t_i)}$$

$$= \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i^2 - 2\bar{t} t_i + \bar{t}^2) + \Sigma (\bar{t} t_i - \bar{t}^2)} = \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i - \bar{t})^2 + \Sigma (\bar{t} t_i) - \bar{t}^2} = \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i - \bar{t})^2 + \bar{t}^2 - \bar{t}^2}$$

$$\alpha = \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i - \bar{t})^2}$$

<u>or</u>

$$\alpha = \frac{\frac{1}{m}\Sigma t_i b_i - \bar{b}\bar{t}}{\frac{1}{m}\Sigma t_i^2 - \bar{t}^2} = \frac{\Sigma t_i b_i - m\bar{b}\bar{t}}{\Sigma t_i^2 - m\bar{t}^2}$$

$$\beta = \bar{b} - \alpha\bar{t} = \bar{b} - \frac{\bar{t}\Sigma t_i b_i - m\bar{b}\bar{t}^2}{\Sigma t_i^2 - m\bar{t}^2} = \frac{\bar{b}\Sigma t_i^2 - m\bar{b}\bar{t}^2 - \bar{t}\Sigma t_i b_i + m\bar{b}\bar{t}^2}{\Sigma t_i^2 - m\bar{t}^2}$$

$$\beta = \frac{\bar{b}\Sigma t_i^2 - \bar{t}\Sigma t_i b_i}{\Sigma t_i^2 - m\bar{t}^2}$$

$$\boxed{\begin{array}{c} \text{The line of best fit is } y = \alpha t + \beta \text{ with} \\ \alpha = \dfrac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i - \bar{t})^2} = \dfrac{\Sigma t_i b_i - m\bar{b}\bar{t}}{\Sigma t_i^2 - m\bar{t}^2} \qquad \beta = \dfrac{\bar{b}\Sigma t_i^2 - \bar{t}\Sigma t_i b_i}{\Sigma t_i^2 - m\bar{t}^2} \end{array}}$$

<u>Alternatively</u> we could write this using $\beta = \bar{b} - \alpha\bar{t}$ as

$$\boxed{y = \alpha t + \bar{b} - \alpha\bar{t} = \bar{b} + \alpha(t - \bar{t}) \text{ with } \alpha = \frac{\Sigma (t_i - \bar{t}) b_i}{\Sigma (t_i - \bar{t})^2} = \frac{\Sigma t_i b_i - m\bar{b}\bar{t}}{\Sigma t_i^2 - m\bar{t}^2}}$$