Chapter 5

5.1 Least Square Problems

$Ax = b$, $A_{m\times n}$, $m \geq n$, $\text{rank}(A) = n$, minimize $\|Ax - b\|_2$.

Minimized $\Leftrightarrow$ gradient $= 0$:

$$\begin{aligned}
\|Ax-b\|_2^2 &= (Ax-b)^T(Ax-b) && (5.1)\\
\lim_{e\to 0}\frac{e^TA^TAe}{\|e\|_2} &= \lim_{e\to 0}\frac{\|Ae\|_2^2}{\|e\|_2} \\
&\leq \lim_{e\to 0}\frac{\|A\|^2\|e\|^2}{\|e\|} \\
&\leq \underbrace{\lim_{e\to 0}\|A\|^2\|e\|}_{\to 0} \\
&= 0 && (5.2)\\
0 &= \lim_{e\to 0}\frac{(A(x+e)-b)^T(A(x+e)-b)-(Ax-b)^T(Ax-b)}{\|e\|_2} \\
&= \lim_{e\to 0}\frac{2e^T(A^TAx-A^Tb)}{\|e\|_2} + \underbrace{\lim_{e\to 0}\frac{e^TA^TAe}{\|e\|_2}}_{=0} \\
&= \lim_{e\to 0}\frac{2e^T(A^TAx-A^Tb)}{\|e\|_2} && (5.3)
\end{aligned}$$

For all $e \to 0$:

$$\begin{aligned}
2\frac{e^T}{\|e\|}(A^TAx-A^Tb) &= 0 && (5.4)\\
A^TAx &= A^Tb && (5.5)
\end{aligned}$$

Why does $x = (A^T A)^{-1} A^T b$ minimize $\|Ax - b\|_2^2$?
Let $x' = x + e$

$$\begin{aligned}
(Ax' - b)^T (Ax' - b) &= (Ae + Ax - b)^T (Ae + Ax - b) \\
&= (Ae)^T Ae + (Ax - b)^T (Ax - b) + 2(Ae)^T (Ax - b) \\
&= \|Ae\|_2^2 + \|Ax - b\|_2^2 + 2e^T (A^T Ax - A^T b) \\
&= \|Ae\|_2^2 + \|Ax - b\|_2^2 \qquad (5.6)
\end{aligned}$$

Minimized for $Ae = 0$.
In practice we never solve $A^T Ax = A^T b$ (called normal equations), instead we solve:

$$A = QR \qquad (5.7)$$

$$(QR)^T (QR)x = (QR)^T b \qquad (5.8)$$

$$R^T Rx = R^T Q^T b \qquad (5.9)$$

$$Rx = Q^T b \qquad (5.10)$$

$$x = R^{-1} Q^T b \qquad (5.11)$$