Warum minimiert $(S^TS)^{-1}Sy$ die Flkt. $\|Sx-y\|_2^2$?

Notiztitel 19.03.2014

1.) In der Lösung war behauptet dass das x für das $\|Sx-y\|_2^2$ minimal ist, der Gleichung $S^TSx = S^Ty$ genügt.

2.) Angenommen die Spalten von S sind lin. unabhängig

$\rightsquigarrow$ $S^T S$ ist invertierbar (warum?)

$\rightsquigarrow$ Das optimale x ist gegeben durch

$$x = (S^T S)^{-1} S^T y$$

3.) Warum ist das so?

Nehmen wir an dass $\underline{S}$ (wie oben) aus lin. unabh. Spaltenvektoren besteht $\rightsquigarrow$ Die orthogonale Projektion auf den von den Spalten-vektoren erzeugten Unterraum $\underline{U}$

$$P_U(y) = S(S^TS)^{-1}S^Ty$$

Aus der Aufgabe 8 wissen wir dass für jeden Vektor u aus U gilt

$$\| P_U(y) - y \|^2 \leq \| u - y \|_2^2$$

Aus der Tatsache daß

$$P_U(y) = Sx \quad \text{für} \quad x = (S^TS)^{-1}S^Ty$$

& der Tatsache dass jeder Vektor u aus U sich darstellen lässt als $u = Sv$ (für ein geeignetes v) folgt

$$\boxed{x = (S^TS)^{-1}S^Ty \quad \Rightarrow \quad \| Sx - y \|_2^2 \leq \| Sv - y \|_2^2 \quad v \text{ bel.}}$$

Tipp: Überprüfen Sie dass Sie den obigen Ansatz verstanden haben indem Sie den Fall, dass S nicht ein unabh. Spaltenvektoren hat behandeln.

Tipp: Denken Sie dabei dass ein Ausdruck $x = (S^T S)^{-1} y$ die verkürzte Form von "x ist die Lösung von $S^T S x = y$"

- Kann man die durch eine ersetzen?
- Denken Sie an das Beispiel $S =$ Nullmatrix.

Alternative Argumentation: $\|Sx-y\|^2 = f(x)$ verhält sich wie eine Quadratische Funktion!

$\rightsquigarrow$ wir untersuchen
$$\|S(x+h)-y\|^2 = \|Sx-y\|^2 + 2\langle Sx-y, Sh\rangle + \|Sh\|^2$$

Ist $0 = S^T(Sx-y)$, dann verschwindet das mittlere Skalarprodukt und wir erhalten mit $z = x+h$

$$\|Sz-y\|^2 = \|Sx-y\|^2 + \|Sh\|^2 \geq \|Sx-y\|^2$$

$\rightsquigarrow$ Bel.