Generalizing the Condition number to Full Rank Least Squares

We were considering the perturbed full rank LS problem

$$\min_x \|(A+tE)x(t)-(b+te)\|_2,$$

where t is real, and E and e are arbitrary but fixed. The normal equations

$$(A+tE)^t(A+tE)x(t) = (A+tE)^t(b+te) \tag{1}$$

were differentiated, giving

$$\dot{x} \equiv \dot{x}(0) = (A^tA)^{-1}[A^t(e-Ex_{LS}) + E^t(b-Ax_{LS})]. \tag{2}$$

Now let's fill in some details from the previous page. Use $x \equiv x(0) \equiv x_{LS}$, and take $\|\cdot\| \equiv \|\cdot\|_2$ throughout (so in particular (i) $\|M\| = \|M^t\|$, (ii) orthogonal matrices are isometries, and (iii) for all matrix actions $\|MN\| \le \|M\|\|N\|$).

If $A = U\Sigma V^t$ is a singular value decomposition, then $\|A\| = \|\Sigma\| = \sigma_1$, (since V and U are orthogonal matrices), and

$$\|(A^tA)^{-1}A^t\| = \|V(\Sigma^t\Sigma)^{-2}V^tV\Sigma U^t\| = \|V(\Sigma^t\Sigma)^{-2}\Sigma U^t\| = 1/\sigma_n,$$

giving the identities (you should verify)

$$\kappa_2(A) \equiv \sigma_1/\sigma_n = \|A\|\|(A^tA)^{-1}A^t\| = \sqrt{\|(A^tA)^{-1}\|\|A\|^2}.$$

The inequality we were working with (a linear Taylor approx to $\|x(t)-x\|/\|x\|$) was

$$|t|\frac{\|\dot{x}(0)\|}{\|x\|} \le |t|\left[\frac{\|(A^tA)^{-1}A^t(e-Ex_{LS})\|}{\|x\|} + \frac{\|(A^tA)^{-1}E^tr\|}{\|x\|}\right] + \mathrm{O}(t^2).$$

Let's break it into the 3 terms (using triangle inequality):

$$|t|\frac{\|\dot{x}(0)\|}{\|x\|} \le |t|\frac{\|(A^tA)^{-1}A^te\|}{\|x\|} + |t|\frac{\|(A^tA)^{-1}A^tEx\|}{\|x\|} + |t|\frac{\|(A^tA)^{-1}E^tr\|}{\|x\|} + \mathrm{O}(t^2),$$

... and massage a bit (same 3 terms, using $\|MN\| \le \|M\|\|N\|$ and $\|E^t\| = \|E\|$):

$$\frac{\|t\dot{x}(0)\|}{\|x\|} \le \frac{\|(A^tA)^{-1}A^t\|\|A\|\|\|te\|}{\|A\|\|x\|} + \frac{\|(A^tA)^{-1}A^t\|\|A\|\|tE\|\|x\|}{\|A\|\|x\|} + \frac{\|(A^tA)^{-1}\|\|A\|^2\|\|tE\|\|r\|}{\|A\|^2\|x\|} + \mathrm{O}(t^2)$$

$$= \frac{\kappa(A)\|te\|}{\|A\|\|x\|} + \frac{\kappa(A)\|tE\|}{\|A\|} + \frac{\|A\|^2\|(A^tA)^{-1}\|\|tE\|\|\|r\|}{\|A\|\|A\|\|x\|} + \mathrm{O}(t^2).$$

If θ is the angle between b and ColSp(A)), then $s \equiv \sin\theta = \|r\|/\|b\|$. With $c \equiv \cos\theta$, $\|Ax\| = c\,\|b\| \le \|A\|\|x\|$ (draw your projection picture), and we return to the first page (again, same 3 terms until we rearrange on the last line):

$$\frac{\|t\dot{x}(0)\|}{\|x\|} \le \kappa(A)\left[\frac{\|te\|}{c\|b\|} + \frac{\|tE\|}{\|A\|}\right] + \frac{\|A\|^2\|t(A^tA)^{-1}\|\|tE\|\|r\|}{\|A\|c\|b\|} + \mathrm{O}(t^2).$$

$$= \kappa(A)\left[\frac{\|te\|}{c\|b\|} + \frac{\|tE\|}{\|A\|}\right] + \frac{\kappa^2(A)\|tE\|s}{\|A\|c} + \mathrm{O}(t^2).$$

$$= \frac{\|tE\|}{\|A\|}\left(\kappa(A) + \tan(\theta)\,\kappa^2(A)\right) + \frac{\|te\|}{\|b\|}\sec(\theta)\,\kappa(A) + \mathrm{O}(t^2).$$