The xz-Neighborhood to the central region

Jos F. Sturm

May 28, 2003

Given $w \in \Re^n_{++}$, we define a (variable metric) inner product as follows:

$$\langle x, y\rangle_w = \sum_{i=1}^n \frac{x_i y_i}{w_i}.$$

Based on this inner product, we define the (variable metric) norm in the usual way, viz.

$$\|x\|_w := \sqrt{\langle x, x\rangle_w}.$$

The central region in $\Re^n_+$ with parameter $0 \le \theta \le 1/n$ is

$$\mathrm{CR}(\theta) := \{w \in \Re^n | w_{\min} \ge \theta e^{\mathrm{T}} w\}.$$

The w-projection of $w \in \Re^n_{++}$ onto $\mathrm{CR}(\theta)$ is denoted

$$w^\theta := \arg\min_f\{\|f - w\|_w \mid f \in \mathrm{CR}(\theta)\}.$$

It is known from Theorem 2 that $w^\theta > 0$. We have

$$0 = \langle w^\theta, w^\theta - w\rangle_w = \|w^\theta\|_w^2 - \langle w, w^\theta\rangle_w.$$

We let

$$\sigma_w^\theta := \left(\frac{\|w^\theta - w\|_w}{\|w\|_w}\right)^2,$$

so that $\arcsin\sqrt{\sigma_w^\theta}$ is the angle between w and w^θ (or equivalently, the angle between w and $\mathrm{CR}(\theta)$). Using the relation $\sin^2\phi + \cos^2\phi = 1$, we have

$$\left(\frac{\|w^\theta\|_w}{\|w\|_w}\right)^2 + \sigma_w^\theta = 1. \tag{1}$$

We remark that if $w > 0$ approaches the boundary of $\Re^N_+$ then $w^\theta \to 0$ and $\sigma_w^\theta \to 1$. However, since $w^\theta \ne 0$ it holds that $\sigma_w^\theta < 1$ for all $w \in \Re^n_{++}$.

Observe that

$$\|w\|_w^2 = e^{\mathrm{T}} w,\ \|w^\theta\|_w^2 = \langle w, w^\theta\rangle_w = e^{\mathrm{T}} w^\theta.$$

It is obvious that $(w_i^\theta - w_i)^2/w_i \le \|w^\theta - w\|_w^2$. However, using the fact that $\langle w^\theta, w^\theta - w\rangle = 0$, we can strengthen this relation slightly.

Lemma 1 *Let $w \in \Re^n_{++}$ and $i \in \{1, 2, \ldots, n\}$. It holds that*

$$\frac{(w_i^\theta - w_i)^2}{w_i} \le \left(1 - \frac{(w_i^\theta)^2}{w_i \|w^\theta\|_w^2}\right) \|w^\theta - w\|_w^2.$$

Proof. We let $w_{\neg i}, w_{\neg i}^\theta \in \Re^n$ be the vectors with entries

$$(w_{\neg i})_j = w_j, \quad (w_{\neg i}^\theta)_j = w_j^\theta \text{ for } j \in \{1, 2, \ldots, n\} \setminus \{i\},$$

and $(w_{\neg i})_i = (w_{\neg i}^\theta)_i = 0$. Since

$$0 = \langle w^\theta, w^\theta - w \rangle_w = \frac{w_i^\theta (w_i^\theta - w_i)}{w_i} + \langle w_{\neg i}^\theta, w_{\neg i}^\theta - w_{\neg i} \rangle_w,$$

we have

$$\begin{aligned}
\left(\frac{w_i^\theta (w_i^\theta - w_i)}{w_i}\right)^2 &= \left(\langle w_{\neg i}^\theta, w_{\neg i}^\theta - w_{\neg i} \rangle_w\right)^2 && (2)\\
&\le \|w_{\neg i}^\theta\|_w^2 \|w_{\neg i}^\theta - w_{\neg i}\|_w^2 && (3)\\
&= \|w_{\neg i}^\theta\|_w^2 (\|w^\theta - w\|_w^2 - \frac{(w_i^\theta - w_i)^2}{w_i}), && (4)
\end{aligned}$$

where the inequality is due to Cauchy-Schwarz. Rearranging terms, we find that

$$\left(\frac{(w_i^\theta)^2}{w_i} + \|w_{\neg i}^\theta\|_w^2\right) \frac{(w_i^\theta - w_i)^2}{w_i} \le \|w_{\neg i}^\theta\|_w^2 \|w^\theta - w\|_w^2. \quad (5)$$

However,

$$\frac{(w_i^\theta)^2}{w_i} + \|w_{\neg i}^\theta\|_w^2 = \|w^\theta\|_w^2,$$

and $\|w^\theta\|_w^2 = (1 - \sigma_w)\|w\|_w^2 > 0$ so that

$$\frac{(w_i^\theta - w_i)^2}{w_i} \le \left(1 - \frac{(w_i^\theta)^2}{w_i \|w^\theta\|_w^2}\right) \|w^\theta - w\|_w^2. \quad (6)$$

Q.E.D.

Lemma 2 *Let*

$$\eta_i := \frac{\|w\|_w^2 - w_i}{2 w_i^\theta} \sigma_w^\theta.$$

It holds that

$$1 + \eta_i - \sqrt{(1 + \eta_i)^2 - 1} \le \frac{w_i^\theta}{(1 - \sigma_w^\theta) w_i} \le 1 + \eta_i + \sqrt{(1 + \eta_i)^2 - 1}.$$

or equivalently,

$$1 + \eta_i - \sqrt{(1 + \eta_i)^2 - 1} \le \frac{(1 - \sigma_w^\theta) w_i}{w_i^\theta} \le 1 + \eta_i + \sqrt{(1 + \eta_i)^2 - 1}.$$

Proof. In order to simplify notations, we define

$$\alpha_i := \frac{w_i^\theta}{(1-\sigma_w^\theta)w_i}.$$

Using Lemma 1 and (1), we have

$$\begin{aligned}
((1-\sigma_w^\theta)\alpha_i - 1)^2 &= \frac{(w_i^\theta - w_i)^2}{w_i^2} \\
&\leq \frac{1}{w_i}\left(1 - \frac{(w_i^\theta)^2}{w_i\|w^\theta\|_w^2}\right)\|w^\theta - w\|_w^2 \\
&= \left(\frac{1}{w_i} - \frac{(1-\sigma_w^\theta)^2\alpha_i^2}{\|w^\theta\|_w^2}\right)\sigma_w^\theta\|w\|_w^2 \\
&= \frac{\sigma_w^\theta\|w\|_w^2}{w_i} - \sigma_w^\theta(1-\sigma_w^\theta)\alpha_i^2.
\end{aligned}$$

Rearranging terms yields that

$$(1-\sigma_w^\theta)\alpha_i^2 - 2(1-\sigma_w^\theta)\alpha_i + 1 - \sigma_w^\theta + \left(1 - \frac{\|w\|_w^2}{w_i}\right)\sigma_w^\theta \leq 0.$$

By dividing both sides by $1-\sigma_w^\theta > 0$ we obtain that

$$\begin{aligned}
0 &\geq \alpha_i^2 - 2\alpha_i + 1 + \frac{w_i - \|w\|_w^2}{(1-\sigma_w^\theta)w_i}\sigma_w^\theta \\
&= \alpha_i^2 - 2\alpha_i + 1 - 2\eta_i\alpha_i \\
&= (1+\eta_i-\alpha_i)^2 - (2\eta_i + \eta_i^2).
\end{aligned}$$

This concludes the proof. **Q.E.D.**

Proximity measure. We define a proximity measure δ_w^θ as follows:

$$\delta_w^\theta = \frac{\sigma_w^\theta}{4\theta(1-\sigma_w^\theta)} = \frac{\|w^\theta - w\|_w^2}{4\theta\|w^\theta\|_w^2}. \tag{7}$$

Since for all i,

$$\eta_i = \frac{\|w\|_w^2 - w_i}{2w_i^\theta}\sigma_w^\theta \leq \frac{\|w\|_w^2}{2w_i^\theta}\sigma_w^\theta = \frac{\|w^\theta\|_w^2}{2(1-\sigma_w^\theta)w_i^\theta}\sigma_w^\theta \leq \frac{\sigma_w^\theta}{2(1-\sigma_w^\theta)\theta} = 2\delta_w^\theta, \tag{8}$$

we arrive at the following corollary. The corollary provides for any w an explicited central region that contains w and is strictly contained in $\Re_{++}^n \cup \{0\}$.

Corollary 1 *Let $w \in \Re_{++}^n$. Then*

$$w \in \text{ CR}\left(\left(1 + 2\delta_w^\theta - \sqrt{(1+2\delta_w^\theta)^2 - 1}\right)\theta\right).$$

1 Centering Step

The centering step is defined by

$$Z\Delta x + X\Delta z = w^\theta - w. \tag{9}$$

We have that

$$w(1) = w^\theta + \Delta X \Delta z. \tag{10}$$

Recall now the well known (see MTY) relations

$$\|\Delta X \Delta z\|_\infty \le \frac{\|Z\Delta x + X\Delta z\|_w^2}{4}, \; \|\Delta X \Delta z\|_2^2 \le \frac{\|Z\Delta x + X\Delta z\|_w^4}{8}, \tag{11}$$

which, together with (9) yield

$$\|\Delta X \Delta z\|_\infty \le \theta \delta_w^\theta \|w^\theta\|_w^2 \|\Delta X \Delta z\|_2^2 \le 2(\theta\delta_w^\theta)^2 \|w^\theta\|_w^4. \tag{12}$$

Lemma 3 *Suppose that*

$$\delta_w^\theta < 1. \tag{13}$$

For the centering step as defined in (9), it holds that $w(t) > 0$ *for all* $0 \le t \le 1$. *Furthermore,*

$$\delta_{w(1)}^\theta \le \frac{(\delta_w^\theta)^2}{2(1-\delta_w^\theta)}.$$

Proof. Using (11) we have

$$\|\Delta X \Delta z\|_\infty \le \frac{\|w^\theta - w\|_w^2}{4} = \frac{\sigma_w^\theta \|w\|_w^2}{4} = \frac{1}{4} \frac{\sigma_w^\theta}{1-\sigma_w^\theta} \|w^\theta\|_w^2.$$

We have that

$$w_i(t) = (1-t)w_i + t w_i^\theta + t^2 \Delta x_i \Delta z_i \ge (1-t)w_i + t(w_i^\theta - t\|\Delta X \Delta z\|_\infty). \tag{14}$$

For $0 \le t \le 1$ we have that

$$w_i^\theta - t\|\Delta X \Delta z\|_\infty \ge \left(\theta - \frac{t}{4}\frac{\sigma_w^\theta}{1-\sigma_w^\theta}\right)\|w^\theta\|_w^2 \tag{15}$$

It follows from (14)–(15) and condition (13) that $w(t) > 0$ if $0 \le t \le 1$. Furthermore,

$$\min_i w_i(1) \ge \left(\theta - \frac{1}{4}\frac{\sigma_w^\theta}{1-\sigma_w^\theta}\right)\|w^\theta\|_w^2 = (1-\delta_w^\theta)\theta\|w^\theta\|_w^2. \tag{16}$$

We have

$$\begin{aligned}
\sigma_{w(1)}^\theta \|w(1)\|_{w(1)}^2 &= \|w(1)^\theta - w(1)\|_{w(1)}^2 \\
&\le \|(1-\sigma_{w(1)}^\theta)w^\theta - w(1)\|_{w(1)}^2 \\
&= \|(1-\sigma_{w(1)}^\theta)(w^\theta - w(1)) - \sigma_{w(1)}^\theta w(1)\|_{w(1)}^2 \\
&= \|(1-\sigma_{w(1)}^\theta)\Delta X \Delta z + \sigma_{w(1)}^\theta w(1)\|_{w(1)}^2 \\
&= (1-\sigma_{w(1)}^\theta)^2 \|\Delta X \Delta z\|_{w(1)}^2 + (\sigma_{w(1)}^\theta)^2 \|w(1)\|_{w(1)}^2,
\end{aligned}$$

where we used (10) and the fact that $(\delta x)^{\mathrm{T}}\Delta z = 0$. It is also clear from (10) that

$$\|w(1)\|^2_{w(1)} = \|w^\theta\|^2_w = (1-\sigma^\theta_w)\|w\|^2_w.$$

Hence,

$$\sigma^\theta_{w(1)} \le (1-\sigma^\theta_{w(1)})^2 \frac{\|\Delta X \Delta z\|^2_{w(1)}}{\|w^\theta\|^2_w} + (\sigma^\theta_{w(1)})^2. \tag{17}$$

Denoting

$$\alpha := \frac{\|\Delta X \Delta z\|^2_{w(1)}}{\|w^\theta\|^2_w},$$

we have

$$0 \le (1-\sigma^\theta_{w(1)})^2\alpha - (1-\sigma^\theta_{w(1)})\sigma^\theta_{w(1)} = (1-\sigma^\theta_{w(1)})(\alpha - (1+\alpha)\sigma^\theta_{w(1)}).$$

Since $\sigma^\theta_w < 1$ for all $w > 0$, it follows that

$$\sigma^\theta_{w(1)} \le \frac{\alpha}{1+\alpha}, \tag{18}$$

and therefore

$$\delta^\theta_{w(1)} \le \frac{\alpha}{4\theta}, \tag{19}$$

Using (16) and (12) we can estimate α as follows:

$$\alpha = \frac{\|\Delta X \Delta z\|^2_{w(1)}}{\|w^\theta\|^2_w} \le \frac{\|\Delta X \Delta z\|^2_2}{\|w^\theta\|^2_w(\min_i w_i(1))} \le \frac{2\theta(\delta^\theta_w)^2}{1-\delta^\theta_w}$$

Together with (19) this yields

$$\delta^\theta_{w(1)} \le \frac{(\delta^\theta_w)^2}{2(1-\delta^\theta_w)}.$$

Q.E.D.

2 Affine Scaling Step

The affine scaling step is defined by

$$Z\Delta x + X\Delta z = -w. \tag{20}$$

Furthermore, if $\Delta X \Delta z \ne 0$ we let

$$1/\alpha^* = -\min_i \frac{\Delta x_i \Delta z_i}{w_i}, \tag{21}$$

and $\alpha^* = \infty$ otherwise. We let

$$\alpha = \frac{t^2}{1-t},\ w(\alpha) = w + \alpha \Delta X \Delta z$$

so that

$$(X + t\Delta x)(Z + t\Delta z) = (1 - t)w(\alpha). \tag{22}$$

We have

$$\delta^\theta_{w(\alpha)} = \delta^\theta_w \text{ for all } \alpha \text{ if } \alpha^* = \infty.$$

Assume now that $\alpha^* < \infty$. Define

$$\phi := 1 + 2\delta^\theta_w + \sqrt{(1 + 2\delta^\theta_w)^2 - 1},$$

$$\hat{\alpha} := \frac{8\theta(\beta - \delta^\theta_w)\|w\|^2_w}{\phi^2\|\Delta X \Delta z\|_1}.$$

Using the Cauchy-Schwarz inequality, we have

$$\|\Delta X \Delta z\|_1 \leq \|Z\Delta x\|_w \|X\Delta z\|_w = \frac{1}{2}(\|Z\Delta x + X\Delta z\|^2_w - (\|Z\Delta x\|_w - \|X\Delta z\|_w)^2 \leq \frac{\|w\|^2_w}{2}.$$

Hence,

$$1/\hat{\alpha} \leq \frac{\phi^2}{16\theta(\beta - \delta^\theta_w)}. \tag{23}$$

Also,

$$1/\alpha^* = -\min_i \frac{\Delta x_i \Delta z_i}{w_i} \leq \left(\max_i \frac{\|w\|^2_w}{w_i}\right) \frac{\|\Delta X \Delta z\|_\infty}{\|w\|^2_w} \leq \frac{\phi}{4\theta}. \tag{24}$$

where we used (11) and Corollary 1.

Theorem 1 *We have that*

$$\delta^\theta_{w(t)} \leq \beta$$

for all

$$0 \leq t \leq \frac{2}{1 + \sqrt{1 + (4/\alpha^*) + (4/\hat{\alpha})}}.$$

Proof. Observe that

$$\|w(\alpha\|^2_{w(\alpha)} = e^{\mathrm{T}} w(\alpha) = e^{\mathrm{T}}(w + \alpha \Delta X \Delta z) = e^{\mathrm{T}} w = \|w\|^2_w.$$

Define

$$\tilde{\sigma}(\alpha) := \frac{\min_\mu \|\mu w^\theta - w(\alpha)\|^2_{w(\alpha)}}{\|w\|^2_2},$$

i.e. μw^θ is the projection of $w(\alpha)$ onto the multiples of w^θ. Therefore, we can find μ explicitly from

$$0 = \langle w^\theta, \mu w^\theta - w(\alpha)\rangle_{w(\alpha)} \tag{25}$$

yielding

$$\mu = \frac{1 - \tilde{\sigma}(\alpha)}{1 - \sigma^\theta_w}. \tag{26}$$

Obviously, $\sigma^\theta_{w(\alpha)} \le \tilde\sigma(\alpha)$ and

$$\delta^\theta_{w(\alpha)} = \frac{\sigma^\theta_w}{4\theta(1-\sigma^\theta_w)} \le \frac{\tilde\sigma(\alpha)}{4\theta(1-\tilde\sigma(\alpha))}. \tag{27}$$

If $\theta = 1$ then the above inequality becomes an identity. We have

$$\frac{\tilde\sigma(\alpha)}{1-\tilde\sigma(\alpha)} = \frac{1}{1-\tilde\sigma(\alpha)} - 1 == \frac{\sigma^\theta_w}{1-\sigma^\theta_w} + \left(\frac{1}{1-\tilde\sigma(\alpha)} - \frac{1}{1-\sigma^\theta_w}\right). \tag{28}$$

From (25)–(26) it follows that

$$\begin{aligned}
0 &= \langle w^\theta, \frac{1}{1-\sigma^\theta_w} w^\theta - \frac{1}{1-\tilde\sigma(\alpha)} w(\alpha)\rangle_{w(\alpha)} \\
&= \frac{\|w^\theta\|^2_{w(\alpha)}}{1-\sigma^\theta_w} - \frac{\|w^\theta\|^2_w}{1-\tilde\sigma(\alpha)} \\
&= \frac{1}{1-\sigma^\theta_w}\left(\sum_{i=1}^n (w^\theta_i)^2 (\frac{1}{w_i(\alpha)} - \frac{1}{w_i})\right) \\
&\quad - \left(\frac{1}{1-\tilde\sigma(\alpha)} - \frac{1}{1-\sigma^\theta_w}\right)\|w^\theta\|^2_w.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{1}{1-\tilde\sigma(\alpha)} - \frac{1}{1-\sigma^\theta_w} &= \frac{1}{(1-\sigma^\theta_w)\|w^\theta\|^2_w}\sum (w^\theta_i)^2 (\frac{1}{w_i(\alpha)} - \frac{1}{w_i}) \\
&= \frac{-\alpha}{(1-\sigma^\theta_w)^2\|w\|^2_w}\sum \frac{(w^\theta_i)^2 \Delta x_i \Delta z_i}{w_i w_i(\alpha)} \\
&\le \frac{\alpha}{\|w\|^2_w(1-\alpha/\alpha^*}\left(\max_i \frac{w^\theta_i}{(1-\sigma^\theta_w)w_i}\right)^2 \sum_i [\Delta x_i \Delta z_i]_- \\
&= \frac{\alpha}{1-\alpha/\alpha^*}\left(\max_i \frac{w^\theta_i}{(1-\sigma^\theta_w)w_i}\right)^2 \frac{\|\Delta X \Delta z\|_1}{2\|w\|^2_w}.
\end{aligned} \tag{29}$$

From Lemma 2 and (8), we have

$$\max_i \frac{w^\theta_i}{(1-\sigma^\theta_w)w_i} \le 1 + 2\delta^\theta_w + \sqrt{(1+2\delta^\theta_w)^2 - 1} =: \phi$$

Summarizing (27)–(29), we have

$$\delta^\theta_{w(\alpha)} \le \delta^\theta_w + \frac{\alpha}{1-\alpha/\alpha^*}\,\frac{\phi^2\|\Delta X \Delta z\|_1}{8\theta\|w\|^2_w} = \delta^\theta_w + \frac{(\beta-\delta^\theta_w)\alpha/\hat\alpha}{1-\alpha/\alpha^*}. \tag{30}$$

We obtain that $\delta^\theta_w \le \beta$ if

$$t \le \frac{2}{1+\sqrt{1+4/\alpha^* + 4/\hat\alpha}}.$$

Q.E.D.

3 Computing the projection

Theorem 2 *Let $w \in \Re^n_{++}$ and define*

$$h^* = \inf\{h \mid \max h, w \in \text{ CR}(\theta)\}.$$

If $h^* = -\infty$ then $w^\theta = w$. Otherwise, let

$$\begin{cases} T = \{i | w_i < h^*\}, \\ m_1 = \sum_{i \in T} \frac{h^* - w_i}{w_i}, \\ \alpha = \frac{m_1 \theta}{1 + m_1 \theta}. \end{cases}$$

Then

$$w_i^\theta = (1 - \alpha) \max\{h^*, w_i\} \text{ for } i = 1, 2, \ldots, n.$$

It follows that $w^\theta > 0$.

Proof. Consider the Lagrangean

$$L(x, \lambda) = \frac{\|x - w\|_w^2}{2} + \lambda^{\mathrm{T}}(\theta(e^{\mathrm{T}} x) e - x),$$

with derivative

$$\nabla_x L(x, \lambda) = W^{-1}(x - w + (e^{\mathrm{T}} \lambda) w - W \lambda),$$

where $\lambda \geq 0$ and $\lambda_j = 0$ if $w_j > h$. Then $w = w^\theta$ satisfies the KKT conditions with

$$\lambda_i = (1 - \alpha) \frac{h - w_i}{w_i} \text{ for } i \in T.$$

Q.E.D. We also remark that

$$\sigma_w^\theta = \alpha - (1 - \alpha) \frac{\sum_{i \in T} (h - w_i)}{\|w\|_w^2}.$$