Nonexpansiveness of Resolvent Operator

Take $(x,u) \in J_{\lambda T}$ (i.e. $u = (I+\lambda T)^{-1}x$) and $(y,v) \in J_{\lambda T}$, then by definition we have

$$u + \lambda T u \ni x, \qquad v + \lambda T v \ni y.$$

Subtracting these yields

$$u - v + \lambda(Tu - Tv) \ni x - y$$

Taking inner product with $u - v$ and using monotonicity of T gives

$$\langle u-v, u-v\rangle + \lambda\langle u-v, Tu-Tv\rangle = \langle u-v, x-y\rangle \tag{1}$$
$$\|u-v\|^2 \le \langle u-v, x-y\rangle \le \|u-v\|\|x-y\| \tag{2}$$
$$\|u-v\| \le \|x-y\| \tag{3}$$
$$\|(I+\lambda T)^{-1}x - (I+\lambda T)^{-1}y\| \le \|x-y\| \tag{4}$$
$$\tag{5}$$

Nonexpansiveness of Cayley Operator

Take again $(x,u) \in J_{\lambda T}$ (i.e. $u = (I+\lambda T)^{-1}x$) and $(y,v) \in J_{\lambda T}$,

$$\begin{aligned}
\|C_{\lambda T}x - C_{\lambda T}y\|^2 &= \|2(u-v) - (x-y)\|^2 && (6)\\
&= 4\|u-v\|^2 - 4\langle x-y, u-v\rangle + \|x-y\|^2 && (7)\\
&\le 4\langle x-y, u-v\rangle - 4\langle x-y, u-v\rangle + \|x-y\|^2 && (8)\\
&= \|x-y\|^2 && (9)
\end{aligned}$$

Convergence of Krasnosel'skii-Mann Iteration

We'll make use of the identity

$$\|(1-\theta)a + \theta b\|^2 = (1-\theta)\|a\|^2 + \theta\|b\|^2 - \theta(1-\theta)\|a-b\|^2,$$

which holds for any $\theta \in \mathbb{R}$, $a, b \in \mathbb{R}^n$.

Let $F = (1-\theta)I + \theta T$ be averaged, where $\theta \in (0,1)$ and T is nonexpansive. Note that T has the same fixed points as F

$$\begin{aligned}
& u^* = Tu^* && (10)\\
\Leftrightarrow\quad & \theta u^* = \theta T u^* && (11)\\
\Leftrightarrow\quad & (1-\theta)u^* + \theta u^* = (1-\theta)u^* + \theta T u^* && (12)\\
\Leftrightarrow\quad & u^* = [(1-\theta)I + \theta T]\, u^* = F u^* && (13)
\end{aligned}$$

We consider the fixed point iteration

$$u^{k+1} = F u^k = (1-\theta)u^k + \theta T u^k.$$

Denote be U the (nonempty) set of fixed points of F and let $u^* \in U$ be a fixed point of F. Then we have

$$\begin{aligned}
\|u^{k+1} - u^*\|^2 &= \|(1-\theta)(u^k - u^*) + \theta(Tu^k - u^*)\|^2 \\
&= (1-\theta)\|u^k - u^*\|^2 + \theta\|Tu^k - u^*\|^2 - \theta(1-\theta)\|Tu^k - u^k\|^2 \\
&= (1-\theta)\|u^k - u^*\|^2 + \theta\|Tu^k - Tu^*\|^2 - \theta(1-\theta)\|Tu^k - u^k\|^2 && (*)\\
&\le (1-\theta)\|u^k - u^*\|^2 + \theta\|u^k - u^*\|^2 - \theta(1-\theta)\|Tu^k - u^k\|^2 \\
&= \|u^k - u^*\|^2 - \theta(1-\theta)\|Tu^k - u^k\|^2
\end{aligned}$$

This shows that the so called Fejèr monotonicity of the fixed point iteration, i.e., the distance to the solution set (which is closed and convex) decreases at each step.

Applying the inequality k times yields

$$\|u^{k+1} - u^*\|^2 \leq \|u^0 - u^*\|^2 - \theta(1-\theta) \sum_{j=0}^{k} \|Tu^j - u^j\|^2$$

and hence

$$\sum_{j=0}^{k} \|Tu^j - u^j\|^2 \leq \frac{\|u^0 - u^*\|^2}{\theta(1-\theta)},$$

which implies that $\|Tu^k - u^k\| \to 0$, for $k \to \infty$.

From that we can also estimate a convergence rate of the fixed-point residual:

$$\min_{j=0\ldots k} \|Tu^j - u^j\|^2 \leq \frac{\|u^0 - u^*\|^2}{(k+1)\theta(1-\theta)},$$

Since the iterates $\{u^k\}_{k=1}^{\infty}$ lie in the compact set (due to the Fejèr monotonicity)

$$\{u^k\}_{k=1}^{\infty} \subset C = \left\{ v \mid \|v - u^*\| \leq \|u^0 - u^*\| \right\},$$

there exists at least one subsequence $\{u^{k_l}\}_{l=1}^{\infty}$ which converges to some point $\widehat{u}$.

Since $Tu^{k_l} - u^{k_l} \to 0$, we also have that $Fu^{k_l} - u^{k_l} = (F-I)u^{k_l} \to 0$. Since $F-I$ is Lipschitz continuous (as T is nonexpansive) and hence continuous, we have that $F\widehat{u} = \widehat{u}$ and hence the subsequence converges to a point in $\widehat{u} \in U$.

As (*) holds for any point from $u^* \in U$, we can apply it the point $\widehat{u}$ our subsequence converges to. We know that for the iterates of the original sequence the distance to this point is monotonically decreasing,

$$\|u^{k+1} - \widehat{u}\| \leq \|u^k - \widehat{u}\|.$$

Since a subsequence $\{u^{k_l}\}_{l=1}^{\infty}$ of $\{u^k\}_{k=1}^{\infty}$ is converging to $\widehat{u}$, and $\|u^k - \widehat{u}\|$ is monotonically decreasing, we have convergence of the entire sequence to $\widehat{u}$.

Positive definiteness of primal-dual step-size matrix

For $\theta = 1$, symmetry immediately follows from the structure of the matrix. Define the following inner products:

$$\langle \mathrm{T}^{-1}u, u\rangle = \|u\|_X$$
$$\langle \Sigma^{-1}p, p\rangle = \|p\|_Y$$

We show positive definiteness directly by using the definition

$$\langle (u,p), M(u,p)\rangle = \langle \mathrm{T}^{-1}u, u\rangle + \langle \Sigma^{-1}p, p\rangle - 2\langle Ku, p\rangle = \|u\|_X^2 + \|p\|_Y^2 - 2\langle Ku, p\rangle > 0$$

We have

$$-2\langle Ku, p\rangle = -2\langle \Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}} \mathrm{T}^{-\frac{1}{2}} u, \Sigma^{-\frac{1}{2}} p\rangle$$

Using Cauchy-Schwarz and Young's inequality $2ab \leq ca^2 + b^2/c$ for any $a, b, c > 0$ we have

$$\begin{aligned} -2\langle Ku, p\rangle &\geq -2\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}} \mathrm{T}^{-\frac{1}{2}} u\| \|\Sigma^{-\frac{1}{2}} p\| \\ &\geq -\left(c\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2 \|u\|_X^2 + \frac{\|p\|_Y^2}{c} \right) \end{aligned}$$

Now since $\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2 < 1$ there exists some $\varepsilon > 0$ with $(1+\varepsilon)^2 \|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2 = 1$.

Then picking $c = 1 + \varepsilon = \frac{1}{\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|} = \frac{1}{(1+\varepsilon)\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2}$ we have

$$\begin{aligned} -2\langle Ku, p\rangle &\geq -\left(\frac{\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2 \|u\|_X^2}{(1+\varepsilon)\|\Sigma^{\frac{1}{2}} K \mathrm{T}^{\frac{1}{2}}\|^2} + \frac{\|p\|_Y^2}{(1+\varepsilon)} \right) \\ &\geq -\left(\frac{1}{1+\varepsilon} \right) (\|u\|_X^2 + \|p\|_Y^2), \end{aligned}$$

Inserting this estimate back, we have

$$\langle u, u\rangle_X + \langle p, p\rangle_Y - 2\langle Ku, p\rangle \geq \|u\|_X^2 + \|p\|_Y^2 - \left(\frac{1}{1+\varepsilon} \right) (\|u\|_X^2 + \|p\|_Y^2) = \frac{\varepsilon}{1+\varepsilon} \left(\|u\|_X^2 + \|p\|_Y^2 \right)$$

Since $\varepsilon > 0$, this term is bigger than 0 and hence M is positive definite.