# SUPPLEMENTAL MATERIAL TO: "ON THE NUMBER OF ITERATIONS FOR THE MATCHING PURSUIT ALGORITHM"

*Fangyao Li*[1], *Christopher M. Triggs*[1], *Bogdan Dumitrescu*[2], *Ciprian Doru Giurcăneanu*[1]

[1]Department of Statistics
University of Auckland
Private Bag 92019, Auckland 1142, New Zealand

[2]Department of Automatic Control and Computers
University Politehnica of Bucharest
313 Spl. Independenţei, 060042 Bucharest, Romania

## 1. PROOFS FOR PROPOSITIONS 3.1-3.3

In this document, we use the same notation as in [1]. Additionally, for an arbitrary matrix $\mathbf{M}$, we have that $\mathrm{Ker}(\mathbf{M}) = \{\mathbf{v} : \mathbf{M}\mathbf{v} = \mathbf{0}\}$.

We need the following technical results:

**Result 1.** *Let $\bar{\mathbf{x}}, \mathbf{y} \in \mathbb{R}^n$ such that $||\bar{\mathbf{x}}|| = 1$ and $\mathbf{y} \neq \mathbf{0}$. With the convention that $\mathbf{P} = \bar{\mathbf{x}}\bar{\mathbf{x}}^\top$, we have:*

$$||(\mathbf{I} - \nu\mathbf{P})\mathbf{y}|| \leq ||\mathbf{y}|| \text{ if } \nu \in (0, 1]. \tag{1}$$

*The equality is achieved if and only if $\bar{\mathbf{x}}^\top\mathbf{y} = 0$.*

*Proof.* The result can be established by observing that $||(\mathbf{I} - \nu\mathbf{P})\mathbf{y}||^2 = \mathbf{y}^\top(\mathbf{I} - \nu\mathbf{P})^2\mathbf{y}$ and the largest eigenvalue of the symmetric matrix $(\mathbf{I} - \nu\mathbf{P})^2$ is equal to one. Then (1) is a consequence of the well-known Rayleigh inequality. □

**Result 2.** *For $\mathbf{y} \in \mathbb{R}^n$, we have that $||\mathbf{A}_m\mathbf{y}|| \leq ||\mathbf{y}||$ when $m \geq 1$.*

*Proof.* Using the notation $\mathbf{A}_0 = \mathbf{I}$, Result 1 implies that $||(\mathbf{I} - \nu\mathbf{P}_{s(j+1)})(\mathbf{A}_j\mathbf{y})|| \leq ||\mathbf{A}_j\mathbf{y}||$ for all $0 \leq j \leq m-1$. This leads straightforwardly to Result 2. □

**Result 3.** *The following identity holds true for $m \geq 1$:* $\mathrm{tr}(\mathbf{B}_{m+1} - \mathbf{B}_m) = \nu\mathrm{tr}(\mathbf{P}_{s(m+1)}\mathbf{A}_m)$.

*Proof.* We can readily write the identities: $\mathrm{tr}(\mathbf{B}_{m+1} - \mathbf{B}_m) = \mathrm{tr}\left[(\mathbf{I} - \mathbf{A}_{m+1}) - (\mathbf{I} - \mathbf{A}_m)\right] = \mathrm{tr}(\mathbf{A}_m - \mathbf{A}_{m+1})$ $= \mathrm{tr}\left[\mathbf{A}_m - (\mathbf{I} - \nu\mathbf{P}_{s(m+1)})\mathbf{A}_m\right] = \nu\mathrm{tr}(\mathbf{P}_{s(m+1)}\mathbf{A}_m)$. □

**Proof of Prop. 3.1:** An important consequence of Result 3 is that, for proving Prop. 3.1, it suffices to demonstrate the inequality $\mathrm{tr}(\mathbf{P}_{s(m+1)}\mathbf{A}_m) \leq 1$. The fact that $\mathrm{rank}(\mathbf{P}_{s(m+1)}) = 1$ implies $\mathrm{rank}(\mathbf{P}_{s(m+1)}\mathbf{A}_m) \leq 1$. Additionally, we have that $(\mathbf{P}_{s(m+1)}\mathbf{A}_m)\bar{\mathbf{x}}_{s(m+1)} = (\bar{\mathbf{x}}_{s(m+1)}\bar{\mathbf{x}}_{s(m+1)}^\top)\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)} = (\bar{\mathbf{x}}_{s(m+1)}^\top\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)})\bar{\mathbf{x}}_{s(m+1)}$, which demonstrates that the only non-zero eigenvalue of $\mathbf{P}_{s(m+1)}\mathbf{A}_m$ is $\bar{\mathbf{x}}_{s(m+1)}^\top\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)}$. Hence, we get:

$$\begin{aligned}
|\mathrm{tr}(\mathbf{P}_{s(m+1)}\mathbf{A}_m)| &= |\bar{\mathbf{x}}_{s(m+1)}^\top\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)}| \\
&\leq ||\bar{\mathbf{x}}_{s(m+1)}||\,||\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)}|| \quad (2)\\
&= ||\mathbf{A}_m\bar{\mathbf{x}}_{s(m+1)}|| \\
&\leq ||\bar{\mathbf{x}}_{s(m+1)}|| = 1. \quad (3)
\end{aligned}$$

The inequality in (2) is obtained by using the properties of the scalar product [2, Th. 1.1], while the inequality in (3) is based on Result 2.

The equality holds in [1, Eq. (5)] if and only if we have simultaneously equalities in (2) and (3). As we know from Result 1 that $||(\mathbf{I} - \nu\mathbf{P}_{s(j)})\bar{\mathbf{x}}_{s(m+1)}|| \leq ||\bar{\mathbf{x}}_{s(m+1)}||$ for any $j \in \{1, \ldots, m\}$, the only possibility for having equality in [1, Eq. (5)] is $\bar{\mathbf{x}}_{s(m+1)} \in \bigcap_{j=1}^m \mathrm{Ker}(\mathbf{P}_{s(j)})$. The condition is equivalent to $\bar{\mathbf{x}}_{s(m+1)}^\top\bar{\mathbf{x}}_{s(j)} = 0$ for all $j \in \{1, \ldots, m\}$.

**Proof of Prop. 3.2:** (i) Using the identity in [1, Eq. (1)], it is easy to show that $\mathbf{B}_m$ is idempotent if and only if $\mathbf{A}_m$ is idempotent. Another important observation is that $\det(\mathbf{A}_m) = (1-\nu)^m$, where $\det(\cdot)$ denotes the determinant. This is a consequence of the fact that $\det(\mathbf{I} - \nu\mathbf{P}_{s(j)}) = 1 - \nu$ for $j \in \{1, \ldots, m\}$. As $\nu \in (0, 1)$, we have $\det(\mathbf{A}_m) \in (0, 1)$. Therefore, $\mathbf{A}_m$ is not idempotent because the determinant of an idempotent matrix can only be zero or one.

(ii) We have from hypothesis that $\mathbf{P}_{s(i)}\mathbf{P}_{s(j)} = \mathbf{0}$ for $m \geq i > j \geq 1$. This property together with the identities in [1, Eqs. (3)-(4)] lead to the conclusion that $\mathbf{A}_m = \mathbf{I} - (\mathbf{P}_{s(m)} + \cdots + \mathbf{P}_{s(1)})$. It follows from [1, Eq. (1)] that $\mathbf{B}_m = \mathbf{P}_{s(m)} + \cdots + \mathbf{P}_{s(1)}$. It is easy to check that $\mathbf{B}_m$ is idempotent and symmetric.

**Proof of Prop. 3.3:** (i) Assume that

$$\begin{aligned}
\mathbf{A}_m^\top\mathbf{A}_m + \mathbf{B}_m^\top\mathbf{B}_m &= \mathbf{I}, \text{ or equivalently,} \quad (4)\\
2\mathbf{A}_m^\top\mathbf{A}_m - \mathbf{A}_m - \mathbf{A}_m^\top &= \mathbf{0}. \quad (5)
\end{aligned}$$

Let $\mathbf{v}$ be an eigenvector of $\mathbf{A}_m$ corresponding to the eigenvalue $\lambda$. Using the fact that $\mathbf{A}_m\mathbf{v} = \lambda\mathbf{v}$ together with (5), we get: (a) $\mathbf{A}_m^\top\mathbf{v} = \frac{\lambda}{2\lambda-1}\mathbf{v}$, which shows that $\frac{\lambda}{2\lambda-1}$ is an eigenvalue for $\mathbf{A}_m^\top$. As the eigenvalues of $\mathbf{A}_m^\top$ are the same with

E-mails: lfan523@aucklanduni.ac.nz, cm.triggs@auckland.ac.nz, bogdan.dumitrescu@acse.pub.ro, c.giurcaneanu@auckland.ac.nz.

| $\varsigma^2$ | $\nu$ | $SC_1$ | $SC_2$ | $ESC_1$ | $ESC_2$ | $gMDL_1$ | $gMDL_2$ | $EgMDL_1$ | $EgMDL_2$ | BIC | EBIC | $AIC_C$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Model 1 | | | | | | |
| 8.0 | 0.95 | 0.1933 | 0.1933 | **0.1745** | **0.1736** | 0.1932 | 0.1933 | **0.1747** | **0.1682** | 0.1941 | 0.1941 | 0.1927 |
| 8.0 | 0.10 | **0.1263** | **0.1263** | **0.1262** | **0.1263** | **0.1263** | **0.1263** | **0.1262** | **0.1262** | **0.1265** | **0.1265** | **0.1259** |
| 0.2 | 0.95 | 1.8986 | 1.8985 | **0.9257** | **0.9170** | 1.8979 | 1.8980 | **0.9417** | **0.9257** | 1.9049 | 1.9049 | 1.8761 |
| 0.2 | 0.10 | **1.1261** | **1.1265** | **1.1255** | **1.1256** | **1.1261** | **1.1261** | **1.1254** | **1.1254** | **1.1279** | **1.1279** | **1.1047** |
| | | | | | | Model 2 | | | | | | |
| 8.0 | 0.95 | 0.0163 | 0.0163 | **0.0132** | **0.0132** | 0.0163 | 0.0163 | **0.0134** | **0.0133** | 0.0163 | 0.0163 | 0.0163 |
| 8.0 | 0.10 | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0113** | **0.0112** |
| 0.2 | 0.95 | 0.0584 | 0.0584 | **0.0266** | **0.0265** | 0.0584 | 0.0584 | **0.0273** | **0.0271** | 0.0586 | 0.0586 | 0.0576 |
| 0.2 | 0.10 | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0362** | **0.0355** |
| | | | | | | Model 3 | | | | | | |
| 8.0 | 0.95 | 1.06 | 1.06 | **0.87** | **0.87** | 1.06 | 1.06 | **0.87** | **0.87** | 1.06 | 1.06 | 1.06 |
| 8.0 | 0.10 | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** | **0.86** |
| 0.2 | 0.95 | 9.61 | 9.61 | **4.50** | **4.50** | 9.61 | 9.61 | **4.55** | **4.55** | 9.65 | 9.65 | 9.48 |
| 0.2 | 0.10 | **5.55** | **5.55** | **5.55** | **5.55** | **5.55** | **5.55** | **5.55** | **5.55** | **5.56** | **5.56** | **5.41** |
| | | | | | | Model 4 | | | | | | |
| 8.0 | 0.95 | 8.10 | 8.10 | **6.31** | **6.28** | 8.10 | 8.10 | **6.38** | **6.37** | 8.10 | 8.10 | 8.08 |
| 8.0 | 0.10 | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.44** | **5.43** |
| 0.2 | 0.95 | 31.51 | 31.51 | **14.94** | **14.94** | 31.51 | 31.51 | **15.08** | **14.94** | 31.58 | 31.58 | 30.94 |
| 0.2 | 0.10 | **18.92** | **18.92** | **18.91** | **18.91** | **18.92** | **18.92** | **18.90** | **18.91** | **18.95** | **18.95** | **18.59** |

**Table 1**. MISE computed for various IT criteria by applying the formula in [1, Eq. (15)] when $n = 20$. For each row of the table, or equivalently for each pair $(\varsigma^2, \nu)$, we show in bold the results which are within a range of $5\%$ from the minimum value on that row.

the eigenvalues of $\mathbf{A}_m$, it follows that $\frac{\lambda}{2\lambda-1}$ is also an eigenvalue for $\mathbf{A}_m$. (b) $(\mathbf{A}_m^\top \mathbf{A}_m)\mathbf{v} = \frac{\lambda^2}{2\lambda-1}\mathbf{v}$, which demonstrates that $\frac{\lambda^2}{2\lambda-1}$ is an eigenvalue for $\mathbf{A}_m^\top \mathbf{A}_m$. Since we know from the proof of Prop. 3.2(i) that $\det(\mathbf{A}_m) = (1-\nu)^m > 0$, we have that the symmetric matrix $\mathbf{A}_m^\top \mathbf{A}_m$ is positive definite. Hence, all the entries of $\mathbf{v}$ are real-valued and $\lambda > 1/2$.

The considerations above imply that the positive numbers $\lambda$ and $\frac{\lambda}{2\lambda-1}$ are eigenvalues for $\mathbf{A}_m$. An important consequence of Result 2 is that both eigenvalues are less than or equal to one. However, if $\lambda \leq 1$, then $\frac{\lambda}{2\lambda-1} \geq 1$. Therefore, we need to have $\lambda = 1$. In other words, all eigenvalues of $\mathbf{A}_m$ are equal to one, which we know it is not possible because $\det(\mathbf{A}_m) = (1-\nu)^m < 1$. We obtained this contradiction because we have assumed that the identity in (4) holds true.

(ii) We have from the proof of Prop. 3.2(ii) that $\mathbf{A}_m$ is idempotent and symmetric, therefore the identity in (5) is true.

## 2. ADDITIONAL RESULTS

For illustrating the case when $p_n > n$ (overcomplete dictionary), we repeat the experiments presented in [1] for $n = 20$. All other experimental settings are the same as in [1, Sec. 5]. The results are reported in Table 1.

## 3. REFERENCES


[1] F. Li, C.M. Triggs, B. Dumitrescu, and C.D. Giurcăneanu, "On the number of iterations for the matching pursuit algorithm," EUSIPCO 2017.

[2] H. Yanai, K. Takeuchi, and Y. Takane, *Projection matrices, generalized inverse matrices, and singular value decomposition*, Springer Science+Business Media, 2011.