ICIC Express Letters
Volume **11**, Number **1**, January **2017**

ICIC International ©2017 ISSN 1881-803X
pp. 11–16

ON THE SOLUTION UNIQUENESS OF THE BASIS PURSUIT MODEL

Zhikai Jiang[1,2] and Hui Zhang[3]

[1]School of Business
[2]School of Science
Huaihai Institute of Technology
No. 59, Cangwu Road, Xinpu Dist., Lianyungang 222005, P. R. China
Zhikai.Jiang@qq.com

[3]College of Science
National University of Defense Technology
Deya Road, Kaifu Dist., Changsha 410073, P. R. China
h.zhang1984@163.com

Received June 2016; accepted September 2016

Abstract. *Conditions to ensure that the basis pursuit (BP) model has a unique solution play a central role for exact sparse recovery. Based on the equivalence between the BP model and the standard linear program, we derive a new necessary and sufficient condition guaranteeing solution uniqueness for the BP model. Moreover, with this condition, we provide an elementary proof to verify an existing uniqueness condition of solution for exact sparse recovery. As a potential application, these conditions stimulate new methods to construct test instances for BP models based on linear programming.*
Keywords: Basis pursuit, Linear programming, Sparse recovery, Uniqueness condition

1. **Introduction.** Recovering an unknown approximately or exactly sparse vector $x^* \in \mathbb{R}^n$ from its linear information Ax^*, where A is an $m \times n$ measurement/sensing matrix with $m \leq n$ typically, is a core problem in many fields, including compressed sensing, signal processing, and statistical learning. One of the most efficient models solving this problem is the well-known BP model

$$\begin{aligned} &\text{minimize} \quad \|x\|_1 \\ &\text{subject to} \quad Ax = Ax^* \end{aligned} \tag{1}$$

which was introduced in [1] and has attracted lots of interest from researchers in the fields mentioned above during the past two decades. There is a large literature on the solution uniqueness conditions, i.e., when the minimizer to the BP model is unique and equal to x^*. In this paper, by the equivalence between the BP model and the standard linear programming model, we will derive a new necessary and sufficient condition for solution uniqueness in the BP model, with which we further provide an elementary proof to verify an existing uniqueness condition of solution in [2].

2. **Main Results.** Before stating our new solution uniqueness result, we introduce the notation used throughout the paper.

For an $m \times n$ matrix A and a subset $I \subset \{1, \cdots, n\}$, A_I denotes the $m \times |I|$ matrix with column indices in I, and I^c is the complement of I. Here $|\cdot|$ denotes the cardinality for a set and the absolute value for a number. Particularly, $A_{\{i\}}$ denotes the i-th column of A, also denoted by a_i for simplicity. Likewise, for a vector $v \in \mathbb{R}^n$, $v_I \in \mathbb{R}^{|I|}$ is the restriction of v to indices in I. The vector e_k will denote a vector of ones of dimension k. For brevity we shall often omit mentioning the dimensionality of a vector or a matrix, when it is obvious from the context. The new solution uniqueness result can be stated as follows.

Theorem 2.1. *Let $I := supp(x^*)$, $I_+ := \{i : x_i^* > 0\}$ and $I_- := \{i : x_i^* < 0\}$. Then x^* is the unique minimizer of the BP model if and only if the following linear system*

$$\left(A_{I_+}, -A_{I_-}\right)^T y = e_k + rp \tag{2a}$$

$$\left(A_{I_+^c}, -A_{I_-^c}\right)^T y \leq e_{2n-k} + rq \tag{2b}$$

has a solution (y, r) with $r \in (0, 1]$ for each (p, q) with $p \in \mathbb{R}^k$, $q \in \mathbb{R}^{2n-k}$.

With the above theorem, we will show that the following condition is also necessary and sufficient for solution uniqueness in the BP model.

Condition 2.1. *Under the definitions $I := supp(x^*) \subseteq \{1, \ldots, n\}$ and $s := \mathrm{sign}(x_I^*)$, matrix $A \in \mathbb{R}^{m\times n}$ has the following properties:*

1) submatrix A_I has full column rank, and
2) there is $d \in \mathbb{R}^m$ obeying $A_I^T d = s$ and $\|A_{I^c}^T d\|_\infty < 1$.

Theorem 2.2. *Let $I := supp(x^*)$. Then x^* is the unique minimizer of the BP model if and only if Condition 2.1 holds.*

Condition 2.1 was first proposed in [2] together with a sufficiency proof for solution uniqueness in the BP model by an approach based on Lagrange dual analysis. [3] considers the BP model with complex-valued quantities and A equal to a down-sampled discrete Fourier operator, for which it establishes both the necessity and sufficiency of Condition 2.1 to the solution uniqueness of the BP model. Their proof uses the Hahn-Banach separation theorem and the Parseval formula. [4] lets the entries of matrix A and vector x in the BP model have complex values and gives a sufficient condition for its solution uniqueness. In the regularization theory, Condition 2.1 is used to derive linear error bounds under the name of range or source conditions in [5], which shows the necessity and sufficiency of Condition 2.1 for solution uniqueness of the BP model in a Hilbert-space setting by a weak null space property and the duality theory. Recent papers [6, 7] show that Condition 2.1 can be applied broadly not only to the BP model but also to the following convex problems

$$\min f_1(Ax - b) + \lambda\|x\|_1 \tag{3}$$

$$\min \|x\|_1, \quad \text{s.t. } f_2(Ax - b) \leq \sigma \tag{4}$$

$$\min f_3(Ax - b), \quad \text{s.t. } \|x\|_1 \leq \tau \tag{5}$$

where λ, σ, $\tau > 0$ are scalar parameters and $f_i(x)$, $i = 1, 2, 3$ are strictly convex functions. The Lasso problem [8] is a special case of problem (3) or (5) while the basis pursuit denoising problem [1] is a special case of problem (4) all with $f_i(\cdot) = \frac{1}{2}\|\cdot\|_2^2$, $i = 1, 2, 3$. By constructing the required dual variable in Condition 2.1 via the simplex method, [6] proves that Condition 2.1 is necessary and sufficient for x^* to be the unique solution of the BP model or the convex models (3)-(5). Test instances for BP models are given in [9] to fulfill Condition 2.1.

In this study, we first prove Theorem 2.1 and then prove the equivalence between the solvability of (2) and Condition 2.1. Theorem 2.2 is, therefore, proved since the solvability of (2) is necessary and sufficient for solution uniqueness, Moreover, by the equivalence between the solvability of (2) and Condition 2.1, the solvability of (2) can also be applied to the convex models (3)-(5). All the arguments in this paper are based only on linear programming knowledge, so the proofs are elementary.

3. **Proof of Theorem 2.1.** We need the following lemma which gives a necessary and sufficient condition for the uniqueness of a solution to standard linear program. This result was proved several times by applying the theorem of alternatives in [10, 11]. Here, we give a direct proof.

Lemma 3.1. *Consider the standard linear program:*

$$\begin{aligned} &\text{minimize} \quad c^T z \\ &\text{subject to} \quad Az = b,\ z \geq 0 \end{aligned} \tag{6}$$

Assume that $b \neq 0$. Let x^* be a solution of (6) and denote $S = \{i : x_i^* = 0\}$. Then x^* is unique if and only if the linear system $Ax = 0$, $c^T x \leq 0$, $x_S \geq 0$, $x \neq 0$ has no solution.

Proof: If the linear system has a solution denoted by $\hat{x}$, it must be nonzero since the constraint $x \neq 0$. Consider $x^* + \hat{x}$. Obviously, $x^* + \hat{x}$ is a feasible point of (6) different from x^* and

$$c^T x^* \leq c^T (x^* + \hat{x}) = c^T x^* + c^T \hat{x} \leq c^T x^*$$

where the first inequality is due to the fact that x^* is a minimizer of (6) and the second inequality follows from the constraint $c^T \hat{x} \leq 0$. Thus, $c^T \hat{x} = 0$. Combining the feasibility of $x^* + \hat{x}$, we conclude that $x^* + \hat{x}$ is also a solution to (6) which contradicts the uniqueness of x^*.

Conversely, if x^* is not the unique solution of (6), we can find a solution for the linear system $Ax = 0$, $c^T x \leq 0$, $x_S \geq 0$, $x \neq 0$. To do this, let $\hat{x}$ be a solution of (6) different from x^* and consider $\hat{x} - x^*$. It remains to check the constraints of the linear system one by one. First, $A(\hat{x} - x^*) = 0$ since $Ax^* = A\hat{x} = b$. Second, $c^T(\hat{x} - x^*) = 0$ since both x^*, $\hat{x}$ are the minimizers of (6). Third, $x^* - \hat{x} \neq 0$ since x^* and $\hat{x}$ are different from each other. Finally, $(\hat{x} - x^*)_S = \hat{x}_S \geq 0$ due to the definition of S and the feasibility of $\hat{x}$. $\square$

Now, we can prove Theorem 2.1.

Proof: [Proof of Theorem 2.1] By letting $\Phi := (A, -A)$ and $z^T := (u^T, v^T)$, we begin with a standard-form linear program which is an equivalent expression of the BP model

$$\begin{aligned} &\text{minimize} \quad e_{2n}^T z \\ &\text{subject to} \quad \Phi z = b,\ z \geq 0 \end{aligned} \tag{7}$$

By "equivalent", we mean that one can obtain solution from each other by rules:

given z^*, obtain $x^* = u^* - v^*$ with $((u^*)^T, (v^*)^T) = (z^*)^T$
given x^*, obtain $(z^*)^T = ((u^*)^T, (v^*)^T)$ with $u^* = \max(x^*, 0)$, $v^* = \max(-x^*, 0)$.

Thus, by Lemma 3.1, x^* is the unique minimizer of the BP model if and only if corresponding linear system

$$\Phi z = 0, \quad -e_{2n}^T z \geq 0, \quad z_H \geq 0, \quad z \neq 0 \tag{8}$$

has no solution, where $H = \{i : z_i^* = 0\}$ with z^* being given by the rules above. We rearrange the order of the columns of Φ and the entries of z based on the index set H and get a rearranged matrix (Φ_H, Φ_{H^c}) and a rearranged vector $(z_H^T, z_{H^c}^T)^T$. Then the linear system (8) can be equivalently described as

$$\Phi_{H^c} z_{H^c} + \Phi_H z_H = 0 \tag{9a}$$

$$-e_{|H^c|}^T z_{H^c} - e_{|H|}^T z_H \geq 0 \tag{9b}$$

$$z_H \geq 0 \tag{9c}$$

$$(z_H^T, z_{H^c}^T)^T \neq 0 \tag{9d}$$

By Mangasarian's stable theorem of the alternative (Theorem 1, [10]), the linear system (9) has no solution if and only if the following alternative linear system

$$\Phi_{H^c}^T y = -e_{|H^c|} + rp \tag{10a}$$

$$\Phi_H^T y \geq -e_{|H|} + rq \tag{10b}$$

$$1 \geq r > 0 \tag{10c}$$

has solution (y, r) for each (p, q). Thus, the solvability of linear system (10) is a sufficient and necessary condition for the BP model having the unique minimizer.

Now, we transform the linear system (10) into another equivalent form. We consider the following rearrangement of Φ based on the index sets I_+ and I_-

$$\Phi = (A, -A) \rightarrow \left(A_{I_+}, A_{I_+^c}, -A_{I_-}, -A_{I_-^c}\right) \tag{11}$$

By the definition of H, since $I = I_+ \bigcup I_-$, we have $|H^c| = |I_+| + |I_-| = |I| = k$, $|H| = 2n - |H^c| = 2n - k$, and

$$\Phi_{H^c} = \left(A_{I_+}, -A_{I_-}\right), \text{ and } \Phi_H = \left(A_{I_+^c}, -A_{I_-^c}\right) \tag{12}$$

With these notations, the solvability of the linear system (10) is equivalent to the following linear system

$$\left(A_{I_+}, -A_{I_-}\right)^T y = -e_k + r_0 p \tag{13a}$$

$$\left(A_{I_+^c}, -A_{I_-^c}\right)^T y \geq -e_{2n-k} + r_0 q \tag{13b}$$

having solution (y, r) with $r \in (0, 1]$ and each (p, q). The difference between linear systems (10) and (13) lies in that we take the sparsity of x^* into consideration. After multiplying a negative sign on the both sides in above linear system and based on the arbitrariness of (p, q), we conclude that the solvability of linear system (2) is equivalent to that of linear system (13), which finishes the proof. $\square$

4. **Proof of Theorem 2.2.** In this section, we will show the equivalence between Condition 2.1 and the solvability of (2); thus, Theorem 2.2 is implied by Theorem 2.1. First, let us show that the solvability of (2) implies Condition 2.1.

Proof: [Proof of the existence of vector d] In order to apply Theorem 2.1 to showing the existence of d in Condition 2.1, we first transform the existence of vector d to the solvability of a linear system. Since $I = I_+ \bigcup I_-$, $A_I^T d = \text{sgn}(x_I^*)$ can be written out more detailedly, that is

$$a_j^T d = 1 \text{ for } j \in I_+ \text{ and } a_j^T d = -1 \text{ for } j \in I_- \tag{14}$$

In a matrix form, it means

$$\left(A_{I_+}, -A_{I_-}\right)^T d = e_{|I|} = e_k \tag{15}$$

For $\|A_{I^c}^T d\|_\infty < 1$, it can be equivalently written as

$$A_{I^c}^T d < e_{|I^c|}, \text{ and } -A_{I^c}^T d < e_{|I^c|} \tag{16}$$

That is

$$\left(A_{I^c}, -A_{I^c}\right)^T d < e_{2|I^c|} \tag{17}$$

Since $I = I_+ \bigcup I_-$, we have $I^c \subset I_+^c$ and $I^c \subset I_-^c$. Thus, we get a stronger linear inequality

$$\left(A_{I_+^c}, -A_{I_-^c}\right)^T d < e_{|I_+^c|+|I_-^c|} = e_{2n-k} \tag{18}$$

in the sense that if d satisfies (18), it also satisfies (17). Therefore, the existence of d in Condition 2.1 can be guaranteed by the solvability of the following linear system

$$\left(A_{I_+}, -A_{I_-}\right)^T d = e_k \tag{19a}$$

$$\left(A_{I_+^c}, -A_{I_-^c}\right)^T d < e_{2n-k} \tag{19b}$$

The latter can be justified by taking $p = 0$, $q > 0$ in the linear system (2). $\square$

Proof: [Proof of the full column-rank of A_I] It suffices to show that $(A_{I_+}, -A_{I_-})$ is full column-rank since (A_{I_+}, A_{I_-}) has the same columns as A_I. By the solvability of (13a), there must exist y_0 such that $(A_{I_+}, -A_{I_-})^T y_0 = e_k$ by taking $p = 0$. Thus, the solvability of (13a) guarantees the solvability of $(A_{I_+}, -A_{I_-})^T y = p$ for each $p \in \mathbb{R}^k$ that immediately implies the full column-rank of $(A_{I_+}, -A_{I_-})$. □

In the following, we will show that Condition 2.1 also implies the solvability of (2).

Proof: If Condition 2.1 holds, then there exists $d \in \mathbb{R}^m$ such that

$$(A_{I_+}, -A_{I_-})^T d = e_k \tag{20a}$$

$$(A_{I_+^c}, -A_{I_-^c})^T d < e_{2n-k} \tag{20b}$$

Since $I = I_+ \bigcup I_-$, it holds $I_+^c = I_- \bigcup I^c$, $I_-^c = I_+ \bigcup I^c$ and $I_+ \bigcap I^c = \emptyset$, $I_- \bigcap I^c = \emptyset$. Thus, we have the following rearrangement

$$\left(A_{I_+^c}, -A_{I_-^c}\right) \rightarrow \left(-A_{I_+}, A_{I_-}, A_{I^c}, -A_{I^c}\right) \tag{21}$$

By the full column-rank of A_I in the first part of Condition 2.1, for each $p \in \mathbb{R}^k$, there is z such that $(A_{I_+}, -A_{I_-})^T z = p$. Now, let us consider $y = d + rz$; we will show this vector is a solution of the linear system (2). In fact, since $(A_{I_+}, -A_{I_-})^T y = e_k + rp$, it suffices to show that for each q there exists $r > 0$ such that $\left(A_{I_+^c}, -A_{I_-^c}\right)^T y \leq e_{2n-k} + rq$, i.e.,

$$\left(A_{I_+^c}, -A_{I_-^c}\right)^T d + r\left(A_{I_+^c}, -A_{I_-^c}\right)^T z \leq e_{2n-k} + rq \tag{22}$$

By the rearrangement (21), we see that $\left(A_{I_+^c}, -A_{I_-^c}\right)^T d$ is composed by

$$(-A_{I_+}, A_{I_-})^T d \text{ and } (A_{I_+^c}, -A_{I_-^c})^T d$$

whose every entries are strictly less than one from (20b). Thus, we have $\left(A_{I_+^c}, -A_{I_-^c}\right)^T d < e_{2n-k}$. Hence, by taking $r > 0$ small enough, inequality (22) holds which completes the proof. □

5. **Conclusion.** In this study, we derive a new necessary and sufficient condition for the solution to the BP model, which is also the condition of exact sparse recovery. Seemingly, the new condition is much more difficult to verify than the existing Condition 2.1 since the former needs to check the solvability of a linear system for infinite pairs (p, q) while the latter only needs to check the solvability of one linear system and the full column-rank of A_I. However, our theoretical analysis shows that both of them are actually equivalent. Another contribution of this paper is an elementary proof for the sufficiency and necessity of Condition 2.1 in guaranteeing that a given solution to the BP model is the unique solution. The arguments in this paper only rely on the linear programming knowledge. Therefore, following the spirit in [9], the solvability of (2) and Condition 2.1 may be applied to constructing test instances for BP models based on linear programming. We leave this as a future work.

Acknowledgment. The work of Zhikai Jiang was supported by the National Science Foundation of China (71401060). The work of Hui Zhang was supported by the Chinese Scholarship Council during his visit to Rice University. Zhikai Jiang and Hui Zhang thank Rice University, CAAM Department, for hosting them.

REFERENCES

[1] S. S. Chen, D. L. Donoho and M. A. Saunders, Atomic decomposition by basis pursuit, *SIAM J. Scientific Computing*, vol.20, pp.33-61, 1999.

[2] J. J. Fuchs, On sparse representations in arbitrary redundant bases, *IEEE Trans. Information Theory*, vol.50, pp.1341-1344, 2004.

[3] E. J. Candès, J. Romberg and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory*, vol.52, pp.489-509, 2006.

[4] J. A. Tropp, Recovery of short, complex linear combinations via l_1 minimization, *IEEE Trans. Inform. Theory*, vol.51, pp.1568-1570, 2005.

[5] M. Grasmair, M. Haltmeier and O. Scherzer, Necessary and sufficient conditions for linear convergence of l_1-regularization, *Comm. Pure Appl. Math.*, vol.64, no.2, pp.161-182, 2011.

[6] H. Zhang, W. T. Yin and L. Z. Cheng, Necessary and sufficient conditions of solution uniqueness in l_1 minimization, *Journal of Optimization Theory and Application*, vol.164, pp.109-122, 2015.

[7] H. Zhang, M. Yan and W. Yin, One condition for solution uniqueness and robustness of both ℓ_1-synthesis and ℓ_1-analysis minimizations, *Advances in Computational Mathematics*, 2016.

[8] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B*, vol.58, no.1, pp.267-288, 1996.

[9] D. Lorenz, Constructing test instances for basis pursuit denoising, *arXiv.org/abs/1103.2897*, 2011.

[10] O. L. Mangasarian, A stable theorem of the alternative: An extension of the Gordan theorem, *Linear Algebra and Its Applications*, vol.41, pp.209-223, 1981.

[11] O. L. Mangasarian, Uniqueness of solution in linear programming, *Linear Algebra and Its Applications*, vol.25, pp.151-162, 1979.