Comprehensive Feature Analysis

Black-Box Optimization - Week 7 Update

Generated: January 27, 2026

Executive Summary

This report presents a comprehensive analysis of feature importance across 8 black-box optimization functions after 7 weeks of data collection. Using 6 different analytical methods (ARD, Correlation Analysis, Permutation Importance, Gradient Boosting, Variance Analysis, and Sensitivity Analysis), we identify critical dimensions for optimization and provide actionable recommendations for Week 8 strategy.

Key Finding	Impact
F2: x1 dominates (84.5%)	CRITICAL: Move by ≤0.002 only
F5: x3 is key (99.6%)	Keep x3=1.0 fixed, tune others
F7: x6 drives output (99.9%)	Increase x6 toward 0.8
F4: Center optimal	All dimensions matter equally
F8: Distributed importance	Use dimensionality reduction

Analysis Methods Overview

Six complementary methods were used to identify important dimensions:

1. ARD (Automatic Relevance Determination)

GP learns length scales per dimension. Small length scale = HIGH importance.

2. Correlation Analysis

Measures linear (Pearson) and monotonic (Spearman) relationships.

3. Permutation Importance

Shuffles each feature and measures performance drop.

4. Gradient Boosting Importance

Tree-based feature importance from GBM splits.

5. Variance Analysis

Identifies which dimensions have been explored most.

6. Sensitivity Analysis

Measures output range when varying each dimension.

Function-by-Function Analysis

F1 (2D) - Sparse Signal Function

Key Findings:
• x2 slightly more important (ARD: 100%, Perm Imp: 52%)
• Both dimensions matter roughly equally
• Signal only appears around [0.45, 0.45] region
• Week 7 moved to [0.48, 0.48] → LOST signal (0.000008)

Week 7 Result: [0.48, 0.48] → 0.000008 (signal lost)
Best Known: [0.45, 0.45] → 0.0128

Action for Week 8:
• Go LOWER than [0.45, 0.45]
• Try [0.42, 0.42] or [0.40, 0.43]
• Systematic grid search in [0.35-0.50] range

F2 (2D) - Ultra-Narrow Peak ■■ CRITICAL

Key Findings:
• x1 DOMINATES (ARD: 84.5%, Perm Imp: 65%)
• Peak at [0.111, 0.100] → 0.1300
• Week 7 moved to [0.11, 0.10] → collapsed to 0.0468 (-64%!)
• **LESSON: 0.001 move = catastrophic loss**

Week 7 Result: [0.11, 0.10] → 0.0468 (disaster!)
Best Known: [0.111, 0.100] → 0.1300

Action for Week 8:
• **Stay at [0.111, 0.100] or move by ≤0.002 only!**
• Use RBF kernel with tiny length scale [0.01, 0.05]
• Pure exploitation mode (maximize mean)
• This is the most critical function - do not take risks!

F3 (3D) - x1 Dominant Function

Key Findings:
• x1 is KEY (ARD: 99.2%, Perm Imp: 60%)
• x2 and x3 basically irrelevant (<1% combined)
• Best: [0.949, 0.966, 0.808] → -0.0786
• Negative function (minimize magnitude)

Week 7 Result: [0.99, 0.99, 0.99] → -0.427 (exploration failed)
Best Known: [0.949, 0.966, 0.808] → -0.0786

Action for Week 8:
• Focus optimization on x1
• Keep x2, x3 around 0.95
• Return near Week 4 best region
• Use Matérn kernel for sharp boundaries

F4 (4D) - Bowl Function (Center Optimal)

Key Findings:
• x3 and x4 matter most (ARD: 53%, 35%)
• All dimensions contribute
• Best at center: [0.5, 0.5, 0.5, 0.5] → -3.986
• Moving away from center = more negative
• 7 weeks confirm: center is global optimum

Week 7 Result: [0.65, 0.65, 0.65, 0.65] → -15.16 (confirmed center better)
Best Known: [0.5, 0.5, 0.5, 0.5] → -3.986

Action for Week 8:
• Stay at [0.5±0.05] for all dimensions
• Damage control mode
• Use Matérn kernel with equal length scales

F5 (4D) - Sharp Peak Function ■ STAR PERFORMER

Key Findings:
• **x3 DOMINATES (ARD: 99.6%!) - THIS IS THE KEY!**
• x1, x2, x4 are secondary (3.4% combined)
• Best: [1.0, 0.853, 1.0, 0.977] → 6158 (Week 7 breakthrough!)
• x3 must be near 1.0 (peak at boundary)
• Small moves → large changes (sharp peak)

Week 7 Result: [1.0, 0.853, 1.0, 0.977] → 6158 (NEW BEST!)
Progress: Week 5: 5549 → Week 6: 5399 → Week 7: 6158

Action for Week 8:
• **Keep x3=1.0 FIXED (do not change!)**
• Micro-tune x1, x2, x4 within ±0.01
• Use RBF with ARD: length_scale=[1.0, 1.0, 0.05, 1.0]
• Local search only - protect this gain!

F6 (5D) - Multi-Dimensional Function

Key Findings:
• Conflicting signals: ARD says x5 (87%), GBM says x3, x4 (35%, 34%)
• Week 7 pushed x5→0.70 → disaster (-1.552)
• Truth: x3, x4 matter most; x5 risky beyond 0.45
• Best: [0.26, 0.18, 0.50, 0.48, 0.41] → -1.092

Week 7 Result: [0.15, 0.15, 0.50, 0.50, 0.70] → -1.552 (x5 push failed)
Best Known: [0.26, 0.18, 0.50, 0.48, 0.41] → -1.092 (Week 5)

Action for Week 8:
• Return to Week 5 values
• Focus micro-optimization on x3 ∈ [0.48-0.52]
• Keep x5 ≤ 0.45 (risky above this!)
• x1, x2 have low variance (under-explored)

F7 (6D) - x6 Dominant Function

Key Findings:
- **x6 is KING (ARD: 99.9%!) - THE KEY DIMENSION!**
- x2 moderately important (some correlation)
- All others negligible (<0.1%)
- Best: [0.038, 0.462, 0.239, 0.171, 0.378, 0.734] → 1.478
- x6 has strong positive correlation (higher x6 = better)

Week 7 Result: [0.038, 0.462, 0.239, 0.171, 0.378, 0.734] → 1.478 (steady progress)
Progress: Week 5: 0.836 → Week 6: 1.435 → Week 7: 1.478

Action for Week 8:
- **Increase x6 toward 0.75-0.80 (positive correlation!)**
- Keep x2 around 0.45-0.50
- Fix other dimensions at current values
- Treat as 2D problem (x2, x6 only)
- Use RBF with ARD: length_scale=[1.0, 0.3, 1.0, 1.0, 1.0, 0.1]

F8 (8D) - High-Dimensional Function

Key Findings:
- Distributed importance (curse of dimensionality)
- x6 identified as most important (ARD: 87%), but others matter too
- All dimensions contribute roughly equally
- Best: [0.177, 0.194, 0.170, 0.194, 0.294, 0.143, 0.109, 0.208] → 9.692
- High variance exploration (Week 4) failed badly (4.180)
- Moderate values (0.1-0.3) perform best

Week 7 Result: [0.177, 0.194, ...] → 9.692 (stable)
Progress: Week 5: 9.643 → Week 6: 9.676 → Week 7: 9.692 (incremental)

Action for Week 8:
- Stay in [0.1-0.3] range for all dimensions
- Use Expected Improvement (balanced exploration/exploitation)
- Consider dimensionality reduction (PCA to 3-4D)
- Ensemble methods for robustness

Feature Importance Summary

Function	Key Dimension(s)	Importance	Week 8 Strategy
F1 (2D)	x1, x2 (equal)	~50% each	Explore lower [0.40-0.45]
F2 (2D)	x1	84.5%	■■ STAY at [0.111, 0.100]!
F3 (3D)	x1	99.2%	Focus on x1, fix others
F4 (4D)	x3, x4	53%, 35%	Stay at center [0.5,0.5,0.5,0.5]
F5 (4D)	x3	99.6%	■ Keep x3=1.0, tune others
F6 (5D)	x3, x4	35%, 34%	Return to Week 5 best
F7 (6D)	x6	99.9%	■ Increase x6 → 0.8
F8 (8D)	Distributed	All ~10-15%	Stay in [0.1-0.3] range

Model Tuning Recommendations

Based on the feature importance analysis, here are specific recommendations for kernel selection and optimization strategy:

Function	Recommended Kernel	Length Scales	Rationale
F1, F2	RBF	[0.05, 0.05]	Narrow peaks, high sensitivity
F3	RBF with ARD	[0.1, 10.0, 10.0]	x1 dominant, constrain others
F4	Matérn (ν=2.5)	[0.2, 0.2, 0.2, 0.2]	Smooth bowl, equal dims
F5	RBF with ARD	[1.0, 1.0, 0.05, 1.0]	x3 tiny scale (key dim)
F6	RBF with ARD	Auto-learn	Focus on x3, x4, x5
F7	RBF with ARD	[1.0, 0.3, 1.0, 1.0, 1.0, 0.1]	x6 small (key), x2 medium
F8	RBF with ARD	Auto-learn	Or use PCA to 3-4D first

Optimization Strategy by Dimension Importance

• **HIGH Importance (ARD > 50%):** Explore ±0.02 around current best, use fine-grained search, trust GP predictions

• **MEDIUM Importance (ARD 20-50%):** Explore ±0.05 around current best, balance exploration/exploitation

• **LOW Importance (ARD < 20%):** Fix at current best values, don't waste queries, reduces effective dimensionality

Acquisition Function Selection

Function	Acquisition	Reason
F1	UCB (κ=2.0)	Optimistic recovery mode
F2	Pure Exploitation	Maximize mean, ultra-narrow peak
F3, F4, F6	Maximize Mean	Least negative (minimize magnitude)
F5	Local Search	Tight bounds, protect gains
F7, F8	Expected Improvement	Balanced exploration/exploitation

Conclusions & Week 8 Priorities

Critical Priorities:

1. F2 Recovery (HIGHEST PRIORITY):
• Stay at [0.111, 0.100] or move by ≤0.002 only
• This is critical - Week 7 lost 64% from 0.001 move!
• Predicted recovery: 0.120-0.130 points

2. F5 Maintenance (PROTECT GAINS):
• Keep x3=1.0 fixed (99.6% importance!)
• Micro-tune x1, x2, x4 within ±0.01
• Current best: 6158 points
• Target: 6150-6200 (maintain or small gain)

3. F7 Improvement (OPPORTUNITY):
• Increase x6 toward 0.75-0.80
• x6 has 99.9% importance with positive correlation
• Current: 1.478, target: 1.50-1.55

Secondary Priorities:

4. F1 Recovery: Explore lower values [0.40-0.45]
5. F3, F4, F6 Damage Control: Return to historical best, minimize losses
6. F8 Maintenance: Stay in [0.1-0.3] range, use Expected Improvement

Expected Week 8 Total:
• Optimistic: 6210 points (+58 from Week 7)
• Realistic: 6190 points (+38 from Week 7)
• Conservative: 6170 points (+18 from Week 7)

Risk Management:
• F2 is critical - do NOT take risks on this function
• F5 provides most points - protect with conservative local search
• F7 offers improvement opportunity - pursue carefully
• Negative functions (F3, F4, F6) - minimize damage only

Appendix A: Methodology Details

ARD (Automatic Relevance Determination):
Gaussian Process kernel learns a separate length scale parameter for each input dimension during maximum likelihood optimization. Smaller length scales indicate the function varies more quickly along that dimension, implying higher importance.

Correlation Analysis:
Pearson correlation measures linear relationships, while Spearman correlation captures monotonic relationships. Both range from -1 to +1, where values close to 0 indicate weak relationships.

Permutation Importance:
Randomly shuffles each feature and measures the decrease in model performance. Features that cause large performance drops when shuffled are more important. Based on 10 repeats for stability.

Gradient Boosting Importance:
Tree-based importance calculated from the weighted sum of splits on each feature across all trees in the ensemble. Higher values indicate features used more frequently for splits.

Variance Analysis:
Measures the variance of input values across all samples. Higher variance indicates a dimension has been explored more thoroughly across different regions.

Sensitivity Analysis:
Measures the range of output values associated with variation in each input dimension. Larger output ranges suggest higher sensitivity to that dimension.